DTU Library

# Simulation of Operators' Response in Emergencies

**Rasmussen, Jens**

*Publication date:*
1986

*Document Version*
Publisher's PDF, also known as Version of record

Link back to DTU Orbit

*Citation (APA):*
Rasmussen, J. (1986). *Simulation of Operators' Response in Emergencies*. Risø National Laboratory. Risø-M No. 2616

DK8700001

Risø-M-2616

RISØ

Risø-M-2616

# Simulation of Operators' Response in Emergencies

Jens Rasmussen

Risø National Laboratory, DK-4000 Roskilde, Denmark
September 1986

RISØ-M-2616

# SIMULATION OF OPERATORS' RESPONSE IN EMERGENCIES

Jens Rasmussen

Abstract. For the simulation of the accidental course of events in industrial process plants, a model is needed of operators' response to the cues presented by the system. A model is proposed, based on the simplifications which can be made when restricting attention to the operator functions having significants for a probabilistic risk analysis, and to only skill and rule based performance, i.e., to responses in the early phase of an accident. The model is based on Brunswik's lens model, a model of the normal task repertoire, and on a taxonomy of human errors.

To bring the model in perspective, a review of the state of the art of cognitive models of human behaviour is included.

Review Project for Joint Research Center Ispra
Contract No. 2529-84-11 ED ISP DK

INIS descriptors: ACCIDENTS, COMPUTERIZED SIMULATION, CONTROL ROOMS, ERRORS, MAN-MACHINE SYSTEMS, REACTOR OPERATORS, RELIABILITY, RISK ASSESSMENT, TAXONOMY

September 1986
Risø National Laboratory, DK-4000 Roskilde, Denmark

ISBN 87-550-1279-5
ISSN 0418-6435

Grafisk Service, Risø 1986

2/3/4

LIST OF CONTENTS

## INTRODUCTION

Development of a model which allows a realistic simulation of operator responses to complex accident scenarios in nuclear power plants is a very ambitious undertaking and will only be possible at present if all the simplifying assumptions are taken into consideration which this specific aim of the simulation makes feasible. The simulation model considered in this report is intended for incorporation in an integrated simulation model which can support a systematic identification of the accidental courses of events which have to be considered in an exhaustive probabilistic risk analysis. One of the fundamental problems in such analyses is at present explicitly to formulate the search strategy applied to define the relevant scenarios, and thus to delimit the coverage (Rasmussen, 1982). One approach to this problem will be to have a simulation model which is capable, in a systematic way, of generating system responses to all relevant component faults and their combinations. For this to be effective in risk management, a well defined coverage determined by explicitly described principles for the search is more important than to seek the widest possible coverage by use of ad hoc expert intuition and imagination. Well defined coverage is important in order to make it possible to decide a posteriori whether a given accidental event was in fact included (Rasmussen and Pedersen, 1984).

For the technical part of an industrial process plant, the DYLAM simulation program (Reina and Amenda, 1981; Amendola, 1984) can systematically identify the accidental courses of events to consider. This is done by defining the relevant fault modes for all components, and by a complete search through all these modes and their combinations to the order defined by the analyst. A similar exhaustive search for all possible erroneous human acts is clearly not possible, and a set of delimiting assumptions and principles have to be chosen. Such principles are proposed in the next section, and are intended to add up to an evolutionary approach in which a model is developed in sequential stages of refinement together with a program for concurrent testing of the assumptions.

## BASIC PRINCIPLES AND ASSUMPTIONS

The requirements from risk analysis. Important simplifications can be drawn from the intended application of identifying the sources of events to include in risk analysis in addition to those having their origin in technical faults. In the first approximation, the model is only intended to identify the relevant scenarios to analyze, not to supply quantitative probability figures. Therefore, failure modes having effects identical with those of technical components need not to be considered; only modes of human errors adding branches to the fault tree already defined by the DYLAM program will have to be considered. As we will see, this in fact means that we do not have to consider simple, separate human errors affecting component performance, only those adding causal coupling between otherwise independent events. Even when quantitative predictions are not intended as in this first approximation, it should, however, be noted that a stop rule for the search for which error modes to include will in fact imply a ranking according to frequencies as well as the consideration of a probability cut-off limit.

Another simplifying principle can be drawn from the present application. The basic requirement is not to have, in the first approximation, a reliable simulation of operator performance, but a simulation generating an envelope including the relevant responses, based on explicitly known principles. The resulting fault tree can be screened for irrelevant responses during the subsequent quantification of probabilities.

Aspects of human adaptation. Another set of simplifying principles can be based on the adaptive features of human behavior. Highly skilled human operators have successfully adapted to the control requirements of the particular plant during all operating conditions normally met. During such circumstances, their behavior will only reflect characteristics of the plant, not of their cognitive mechanisms. As Simon phrased it: "A man, viewed as a behaving system, is quite simple. The apparent complexity of his behavior over time is largely a reflection of the complexity of the environment in which he finds himself. --- A thinking human is an adaptive system --. To the extent he is effectively adaptive, his behavior will reflect characteristics largely of the outer environment (in the light of his goals) and will reveal only a few limiting properties of his inner environment - of the physiological machinery that enables him to think." (Simon 1969, p.25).

A model of successful operator performance during familiar work situations can, therefore, be considered a model of the requirements of the work environment, considering his immediate goals. To make it a model of human performance, it will then be necessary to add the mechanisms that come into play when adaptation breaks down, i.e., a model of relevant human error mechanisms.

A model of human error mechanisms appears to be feasible for the present application where the focus will be on those error modes which will tend to introduce systematic or likely coupling between otherwise independent events or acts. Previous analyses of events including human errors (Rasmussen, 1981; Rasmussen et al., 1981) tend to show that a wide variety of observable errors during unfamiliar work situations can in fact be explained by a fairly low number of psychological mechanisms related to human adaptability (Rasmussen, 1985), and can be seen as the manifestation of interference in performance during infrequent plant disturbances from the vast repertoire of more or less automated routines. This is in good accordance with recent findings in psychological research indicating that fundamental principles such as striving for similarity matching and choosing the path of least effort are underlying the most significant systematic error mechanisms (Reason, 1985a,b). Consequently, a model of the behavior of skilled operators, including a representation of error mechanisms, will as a basic ingredient include a representation of the total task repertoire including in particular the highly automated daily routines.

In this way, the simulation model will serve to identify systematically the triggers for these systematic human error mechanisms which may inadvertently be designed, instructed, or trained into an industrial process system.

From these arguments, it will be realized that the basic ingredients of a model of operator performance will be a model of the total task repertoire representing the "perfect operator" together with a model of the learning and error mechanisms which add the human perspective.

A model of the task repertoire will, consequently, be a major part of the development required. Different approaches to this part of the project seem to be possible, and will require further analysis. The traditional approach will be a classical task analysis, which has, in fact, been made for nuclear power plants. It will, however, be a major task to obtain all the necessary data, in particular since it is not the formal but the actual procedures which are needed, even for daily routines. It may be argued that the collection of descriptions of operator task sequences and related error data is as important for quantitative risk analysis as the collection of data on component specifications and fault modes.

Another approach worth a closer look will be to generate the task repertoire from a systematic analysis of the plant control requirements. To make this approach feasible, a systematic representation of the control requirements in terms of a means-end hierarchy including the basic intentions behind design decisions will be needed. This approach requires a data collection program which would, however, be needed also for the design of advanced operator decision support systems. Derivation from control requirements will identify required task sequences with many degrees of freedom with respect to operator's choice of the order of acts in parts of the tasks, and should be supplemented by heuristic rules related to human adaptation, or derived from interviews of operating staffs.

Finally, the possibility of identifying relevant task information from systematic interrogation of operators should be carefully considered, using recently developed methods for expert knowledge acquisition. Again some basic assumptions about human adaptation may serve to simplify the modelling. The mechanisms underlying errors in terms of interference from frequently encountered tasks or concepts seem to be closely related to the mechanisms of memory access and may be accessible by means of metacognitive judgments when used for systematic interrogation (Reason, 1985b).

Finally, the model should be compatible with the DYLAM simulation concept. Simulation of a process plant is usually structured in a set of differential equations derived from balance considerations related to the system performance at the level of thermodynamic relations. In that case, mapping of physical changes will be rather complicated. Simulation models can also be based on a library of component models representing component functions including a set of component failure functions, as is the case in the JRC DYLAM program. The latter gives a simple mapping of changes caused by component failures, and can systematically generate a complete set of possible event trees. Relevant accident scenarios are found by selection. In this case, the system is decomposed into physical components for which operational states are defined according to fault modes. The possible failed system states are then generated by means of scanning through the standard component states and sorting out the relevant system consequences.

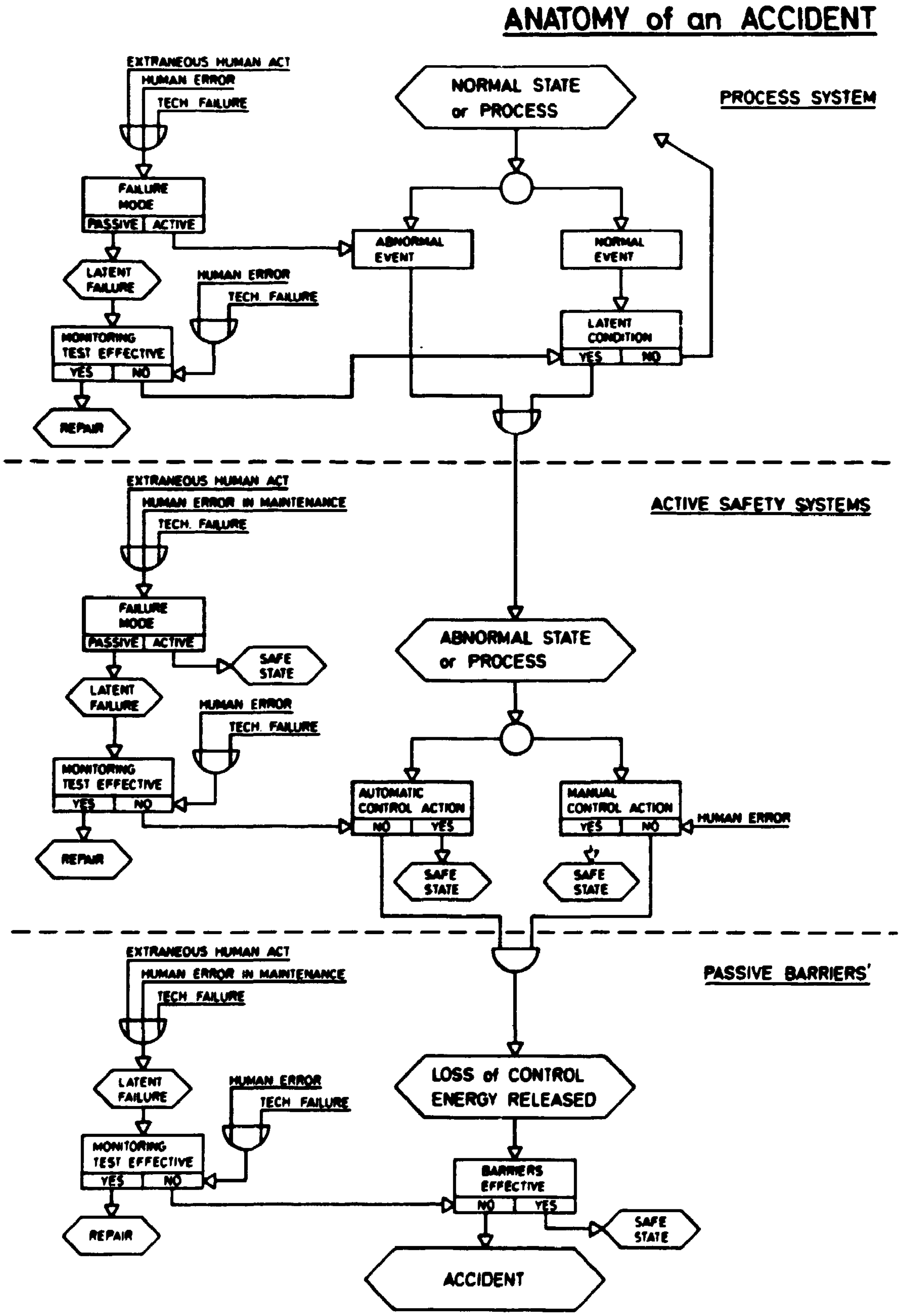



Figure 1. Structure of accidental course of events in a process plant design according to the "defense in depth" philosophy.

In order to do the same for the operator performance, operator behavior should be decomposed into task elements or actions, which may be erroneous due to basic human error modes. In that case, operator performance should be modelled by a representation of the proper task sequence, and erroneous responses generated by a systematic search through a library of relevant error modes. The model approach suggested above, therefore, appears to be immediately applicable in the DYLAM framework.

## SELECTION OF TASKS TO INCLUDE IN THE MODEL

In order to simplify the model of emergency responses, it will be important to formulate the requirements from the present specific application of the model. The operator activities to include in the causal simulation structure can be identified with reference to the anatomy of a typical accident, see Figure 1, and the internal structure of a probabilistic risk analysis (PRA) which are in turn the prerequisites for the DYLAM simulation approach.

The rationale of a PRA for systems with extreme safety requirements such as nuclear installations, depends on the system design concept applied for such plants, leading to a well structured pattern of propagation of accidents. In consequence of the design philosophy applied, the relevant accident scenarios can be grouped and studied collectively, and safety measures can be related to the individual groups by use of a feed back control principle. In addition, accident propagation can be considered to occur in several subsequent phases subject to control by countermeasures based on different physical principles. This "defence-in-depth" design philosophy makes it possible to achieve in a realistic very low probability of an accident with moderate requirements to the probability of the events related to the individual phases of the course of events, <u>given the assumption of mutual independence</u>.

With reference to such general features of an accident sequence, the categories of human activities which should be included in a simulation model can be identified using the consideration that only such activities should be included that will add new branches to the resulting fault and event trees. Activities will not be included if they are merely contributing to the frequency or probability of chains of events which already found from consideration of technical faults. For example, simple errors during human activities in terms of omission of acts or acts that otherwise influence only the reliability of the intended outcome of the task, need not be considered since their consequences will be identical with a lack of response of the technical part involved in the acts. What should in particular be included are human errors which have effects adding coupling and cross links to the physical, causal structure. Such links will introduce coincidence between events related to parts of the physical system which are otherwise functionally independent.

Two categories of such causal links should, in particular, be considered:

One is the systematic source of common mode faults which may be introduced when a basic human error mechanism may lead to operation on a wrong component. In such a case, the lack of a proper act on one component will be coupled to an erroneous act on another. Identification of such links between components which may be functionally unrelated will serve to avoid the combinatorial run-away which will result from purely combinatorial search for relevant common-mode failures.

Typically, this category will result in couplings within one phase of the accident sequence. The category will probably be most important for activities during maintenance, test, and calibration, including activities during work planning and scheduling for such activities, in particular related to redundant protection systems. Characteristically, such activities are of the rule-based domain. The work conditions and, consequently, the related typical error mechanisms will be de-coupled from the dynamics of a particular accidental chain of events. Therefore, a reasonable amount of general empirical evidence will be available, and it will be possible to take account of the influence in fault tree construction, without including the activities in a dynamic simulation.

Another source of causal cross-links is more important for consideration in dynamic simulation, namely the possibility of causal links between the phases of the accidental sequence, i.e., between the initiation of an accidental sequence and the performance of the protective systems. Protective systems are basically feed back control systems serving to maintain the operational state of a system within the acceptable, safe domain. In all systems based on feedback principles, the performance is very sensitive to disturbances of the feed back path. In the present context, operator interaction with the safety systems during accidental chains of events is, therefore, of prime importance. The operator-plant interaction will be very situation specific, and the relevant error mechanisms can only be identified from a joint, dynamic simulation of operator and plant behaviour. Two categories of human activities will be important in this respect. One is the <u>protective tasks</u> which are directly allocated the operating staff, such as control of emergency cooling systems, standby equipment, etc. Another will be the <u>erroneous interference with automatic safety functions</u> due to misunderstanding of the observed functioning. Such interference may be due to errors during unrelated activities caused by, in itself correct, use of a procedure which, however, owing to particular circumstances is counterproductive. The identification of activities to consider can be approached from different points of view. The starting point can be analysis of a task sequence, searching for the effect of human errors (error mode and effect analysis). Another can be search for tasks which may influence a given system component, identified by the fault tree analysis.

In conclusion, for a first approximation, focus should be on modelling the operator interaction with the protective systems of the plant, automatically as well as manually controlled, and only for those error modes that will give rise to additional couplings in the causal structure, in particular across phases in the accidental course of events.

**MODEL OF THE TASK REPERTOIRE**

It follows from the discussion so far that an important part of the model will be a representation of the operator's task repertoire, not only the tasks directly related to the emergency response scenarios, but also all the frequent, normal work routines which may be the source of interference during non-normal situations. Depending upon the search heuristics applied to guide the scanning of error modes, operator tasks will have to be represented in one of two different ways in the data base of the simulation model:

1. Tasks which are directly involved in the course of events that are simulated. These tasks will be selected from the point of view of their possible contribution of new branches to the fault trees which are the ultimate outcome of the DYLAM system. It will not be necessary to include tasks which merely change the frequency, or probability of events in the simulation. Together with the automatic functions, these tasks will be the potential targets of errors and interference.

2. Tasks which can be the potential sources of interference with the protective functions and tasks included in the simulated scenarios. This set includes a representation of the total task repertoire, in particular the "overtrained" daily routines. It should, however, be considered whether the selection of tasks to include can be made by <u>considering only tasks which have action sequence elements or action objects having close relation to the sensitive targets.</u>

For both categories, criteria for the selection of tasks to be included in the model framework are similar to the completeness criteria for risk analysis: It is more important to know the boundaries than it is to have a complete model.

Different approaches to the problem of representing the task repertoire should be considered: The classical approach to task analysis, the systematic derivation from the control requirements of the plant, and interrogation of operating staffs.

**Task analysis**. This approach will involve a considerable amount of work. A recent report (Burgy et al., 1983) describes a major effort to analyze the task repertoire of operating crews. The result has been a computerized data base including 45 operating sequences, with 1062 tasks, including 15,378 task elements. Data are collected from eight nuclear plants, seven simulators, one actual control room, and one mock-up. The data have been collected by operations personnel and human factors specialists, but are limited to observable task characteristics.

The task analysis is based on Miller's (1963) development. A task is defined as the human performance needed to accomplish required system functions independent of the individual who performs it. The entire "single thread" analysis can be considered as a set of links sequentially connected with little, if any, interaction or branching in the link structure.

A controlled vocabulary is used for the analysis. A top-down job function analysis is used to identify and interrelate the detailed task sequences. A task may be performed on different occasions in support of different subfunctions, an operating sequence is a collection of tasks supporting a specific subfunction. Several operating sequences may be performed simultaneously by different crew mebers. Criteria for task selection have been: frequency, criticality, NRC interest, and usability in a "NUREG 0700" review.

Data collection includes analysis of operating procedures, rehearsal of procedure with plant crew to clarify and identify personnel roles, and finally on-site observation.

Task descriptions include:

"Operating sequence overviews" for each sequence to be studied.

"Task sequence charts" identifying the individual elements. The table entries are: sequence no., task and purpose, cue, procedure name and number, and system name.

"Task data forms" identifies: person, location, time, function (behavior element), object of action with state and action, means (cue), and communication link.

A "supplementary task data form" identifies: task difficulty, consequence of error, and criticality of timing; all by qualitative ranking. The data base is extensive and retrieval by computer search is possible, for instance for automatic "link analyses". Unfortunately, however, for the present purpose, the data recorded are not specific enough. The information used to activate task performance ("cues"), for instance, is stated in very general terms, like: "instrument", "procedure", or "verbal". The report will be useful in organizing a suitable data collection, to estimate the amount of resources required, and as a basic structure for collecting the procedures. Therefore, a realistic simulation model of human performance during emergency situations will require a substantial task analysis, not only in terms of formal procedures, but in fact all the normally acceptable variations of procedures in the work repertoire along the lines proposed by Pedersen (1985). This task data collection should, however, also be carefully planned with respect to obtaining task frequency data, etc., when planning the simulation of the "cue scanning strategy" of operators discussed in the human error section below.

In the general context of data banks for risk and reliability analysis, it should be considered whether the data banks including technical component descriptions, their fault modes and frequencies, should also include operator task descriptions in terms of subroutines with activation cues, in order to have a meaningful basis for including systematic human errors in the analysis of response to unfamiliar task requirements.

**Derivation from control requirements.** Since successful task performance in fact reflects successful adaptation to the control requirements of the process plant in question, another possibility of modelling the performance of the "perfect operator" will be the systematic derivation of the control sequences required to meet the operating goals, e.g., to maintain a specified state or to transfer the plant from one specified state to another. Such a systematic control sequence derivation must be based on a consistent representation of the demands and resources in terms of the purpose / function / process / equipment mapping of a means-end hierarchy.

In general, a procedure is a set of rules which describe how actions on the plant should be made if a certain system goal should be accomplished. The sequencing of actions depend on plant structure and functional properties, on the nature of the control task considered, and on the operating constraint, e.g., with respect to safety requirements.

In principle, an operating procedure can be systematically identified by a decomposition of the goal and constraint set of the operating mode considered, top-down through the means-end hierarchy. The result will be a logically consistent specification of a set of concurrent and sequential actions on physical components of the plant. The decomposition will be controlled by the causal topology of the plant at the various levels, and a systematic tool for this derivation can be the multi-level flow model proposed by Lind (1981, 1982), based on a consistent representation of the mass, energy, and information flow topology.

In the operation of a process plant, a distinction can be drawn between two categories of control which are related to different aspects of the coordination of plant functions. These categories are important for the discussion of task structures.

The first category includes the control functions provided for optimization and for maintaining plant integrity during transients caused by external disturbances or by programmed changes in the operating conditions as, e.g., changes of set-points. Characteristic of this type of control functions is that they are provided for a certain operating regime, i.e., they are not applicable if the operating regime is changed. In material and energy processing plants, such as nuclear power plants, this category of control functions performs a coordination of the redistribution of mass and energy stored in plant components, and such coordination problems are related to plant operations where structural changes do not occur.

The second category of controls includes coordination problems related to changes in plant functional structure. This occurs when an integrated process must be established from a set of hitherto functionally unrelated plant components, as for instance during startup, or when establishing ad hoc safety functions. In order to allow process components to be connected, operational conditions for the separate components must be equalized. (A boiler must be filled, heated, and produce steam before it can be connected to the turbine. The turbine must be at correct speed before it can be synchronized with the grid, etc.).

The division of a control task into subtasks according to the categories above leads to a decomposition of the associated goal and procedure into subgoals and subprocedures. Furthermore, to each task a plant subsystem corresponds which again is divided into subsubsystems by the task decomposition. However, plant subsystems obtained in this way will in general be overlapping, i.e., they will share components because the goal decomposition is based on the functional requirements and not on the physical structure. Therefore, a systematic framework is necessary in order to consider all goals and constraints (also related to possible latent fault conditions) in the procedure design, and the application of Lind's approach should be tested. It will be noted that the resulting hierarchical task/goal structure has similarities with the hierarchical task analysis and description presently under development at the JRC-center Ispra (private communication, August 1985), and with the structure proposed by Reason et al. (1985).

The control procedures derived from plant control requirements in this way will have several degrees of freedom in the final sequencing. Functionally speaking, the sequential order in smaller or larger parts of a control sequence will be unspecified and can be chosen from other types of criteria such as minimizing operator motion etc. Operator heuristics and criteria for choice may be important, since there will be a tendency to apply them also in cases where safety considerations dictate task sequences which are "irrelevant" during normal circumstances, but important as conditional protection against potential failure states.

Systematic generation of the control sequence requirements will result in a family of acceptable procedures for each task, and all members of a family should probably be considered as a basis for perturbation according to the error mode search unless clear evidence for operator preferences can be found. In addition, it should be considered to what extent a systematic analysis can identify those parts of a control sequence for which the physical and functional properties of the plant will guarantee error detection by blocking further actions if not corrected. For these aspects of the sequence analysis, Pedersen's (1985), procedures for work analysis should be considered for further development. This approach will also invite a number of questions concerning the necessary data base. A systematic representation of the system properties in terms of the means-end abstraction hierarchy will be necessary. This in turns raises the problem of an explicit formulation of the top-down propagation of the intentional basis of system design. Such information is typically implicitly imbedded in industrial practices, or only present as subjective and unformulated preferences of the designer. An attempt to formulate the information will, however, have general interest, since it is exactly this kind of information which is needed for design of decision support systems for the operators during disturbed system operation, i.e., the information will be needed for the design of "expert systems" intended for on-line decision support.

<u>In consequence, an important study for supplying the data base for systematic generation of the task repertoire, operating procedures, as well as for design of expert system support of operators will be the development of a data base representing the design intentions in terms of a consistent purpose/function/process/ equipment mapping (means-end representation).</u>

<u>Interrogation of operators</u> by more indirect means should be considered a tool for modelling the task repertoire. Recent research (Reason, 1985b) seems to indicate that humans are rather good at metacognitive judgments of frequencies of encounter of concepts, events, objects, etc., and since the predominant feature of task interference seems to be frequency of encounter together with "cue overlap", indirect methods by which operating teams are asked to "spew-out" <u>examplars</u> of categories and to rank frequency of occurrence of cues and task elements may turn out to be a selective and, therefore effective, alternative to the more traditional task analysis. In this respect it will be important to consider the basic organization of performance, discussed in detail in a subsequent section.

Behaviour is composed of sequences of skill-based subroutines, which roll off as integrated smooth units without conscious control of the chaining of the individual acts when activated by the proper intention to act. At the rule-based level, such subroutines are chained by choosing those suitable for the occasion according to know-how or prior experience. If a problem is at hand and no useful task sequence is available, knowledge-based "experiments" by means of a mental model may be necessary to predict the outcome and compare it with relevant goals in order to generate a plan. Only the skill- and rule- based levels will be considered in the interrogation of operating staffs.

Skilled routines are activated as integrated task sequences by an intention or choice, and they will be the elements of the task data base. More complex task sequences controlled at the rule-based level are composed of such routines, which are individually activated by a cue-set including observation of indicators and the result of the antecedent act. An important study will be to ask operators, by use of metacognitive judgement techniques, to list frequencies of the routines and to judge the likelihood of their chaining. According to the findings of Reason (private communication, 1985), well planned questioning techniques should be able to establish antecedent action links, relevant cue sets, and frequency of encounter, i.e., to generate the "frequent-task-interference-source" data base.

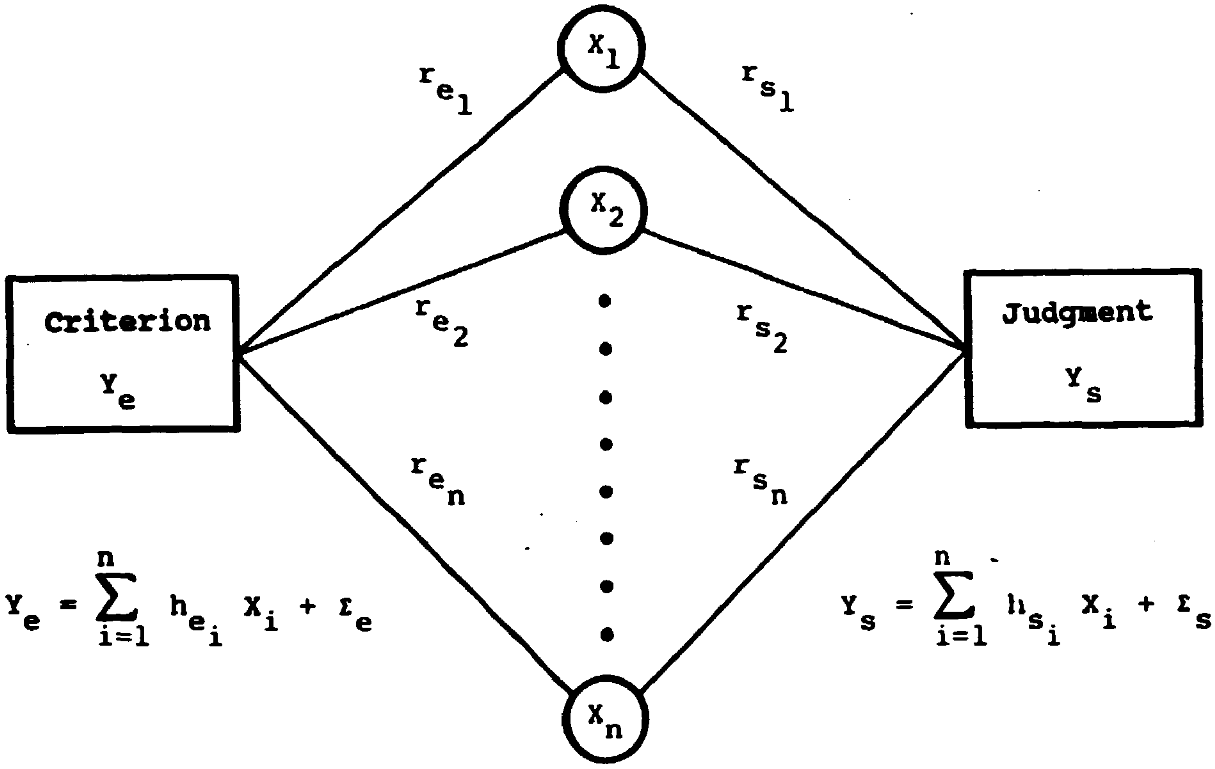
## MODELLING OPERATOR PERFORMANCE

From the introductory discussion it follows that a model of the error free operator represents only task characteristics, and that representation of human limiting properties and error provoking mechanisms is necessary in order to make it a model of human characteristics. The fact that emergency procedures are carefully formulated in advance, and extensively trained on training simulators, together with the dependence of systematic errors on interference with very familiar task situations focus the efforts of the initial modelling upon models at the skill-rule based level. Consulting the review of available model concepts (Appendix 1) it appears that the descriptive framework offered by the Brunswick lens model should be a useful concept.

Arguments for the lens model (figure 2) by Brunswik (1957) are based on the need for judgment because knowledge of the environment is difficult due to "causal ambiguity". Reference is made to Brunswick and Tolman who emphasized that

> "the organism in its intercourse with the environment must cope with **numerous, independent, multiformal relations** among variables which are **partly relevant** and **partly irrelevant** to its purpose, which carry only a **limited amount of dependability**, and which are **organized in a variety of ways.** The problem for the organism, therefore, is to know its environment under these complex circumstances. In the effort to do so, the organism brings a variety of processes (generally labelled cognitive), such as perception, learning, and thinking, to bear on the problem of reducing causal ambiguity. As a part of this effort, human beings often attempt to manipulate variables (by experiments, for example) and sometimes succeed - in such a manner as to eliminate ambiguity. But when the variables cannot be manipulated, human beings must use their cognitive resources unaided by manipulation or experiment. They must do the best they can by passive rather than active means to arrive at a conclusion regarding a state of affairs clouded by causal ambiguity. They must, in short, exercise their judgment. Human judgment is a cognitive activity of last resort." (p.272).

This description appears to be close to the conditions of judgement for process operators. Also the role of the "internal dynamic world model" (Rasmussen, 1986) or "process feel" is in correspondence with the lens model approach. Tolman and Brunswick (1935) react to the notion of information "input". Both argued that the objects and events apprehended by an organism do more -- and less -- than "impinge" upon it. Not only does the organism cognitively act on the "input", but the perceived object carries implications for other objects. That is why Tolman's position was labelled an S-S theory. (sign-significate) and contrasted to S-R theories. Because "cues" and "sign-significates" point outward,




$r_e$ = Ecological correlations

$r_s$ = Response correlations

$h_e, h_s$ = Optimum regression weights

Figure 2. Brunswik's lens model. This model is a descriptive model of human judgement developed within the social judgement paradigm. See Appendix 1 for details.

they involve a relation between two variables - proximal and distal, the given and the inferred. Choice of that relation as the fundamental unit of cognition has profound consequences, of course, and it was this choice which eventually led Tolman to introduce the concept of "mental map" in 1948; he argued that cognition involves a subjective representation of the goal paths in the environment of the organism. Brunswick went further; he demanded a more detailed analysis of the environment and a less detailed analysis of the organism. Environment and organism should be described in symmetrical terms. The lens model therefore considers symmetrically the ecological validity of cues, and the subject's cue utilization.

Central to social judgment theory is the distinction between surface and depth of the environment. surface data are (given) cues to the (inferred) depth conditions in the judgment task. The intervening region between surface and depth has been named zone of ambiguity. The relation is considered a cause (depth) - effect (surface) relation. Because a single effect may be produced by several causes, as well as because multiple effects may be produced by a single cause, there is ambiguity from cause to effect and from effect to cause. Because causes may be interrelated and because effects are interrelated, the network of task relations can be said to be entangled. Moreover, causal ambiguity is produced because (1) surface data are less than perfectly related to depth variables, (2) functional relations between surface and depth variables may assume a variety of forms (linear, curve-linear), and (3) the relations between surface and depth may be organized according to a variety of principles (for example, additive or pattern). These circumstances give more specific meaning to the term "causal texture" or causal "ambiguity".

Objectives for social judgment theory are stated to be:

1. Real life relevance.

2. A descriptive, not a law-seeking theory.

3. Aids to improve human judgment, for instance by displaying pictorially the weights, function forms, and uncertainties in persons' judgment policies as well as in judgment tasks.

The central features of the lens model concept, i.e., the focus on a descriptive model with equal emphasis on the representation of the task environment and the human response, are clearly compatible with the needs in the present context. In some respects, however, reference to the "lens model" concept may be misleading, because the label refers to a methodological paradigm, rather than merely to a model structure, and the similarity to the present approach may be judged a verisimilitude at the present state of the development. However, being a purely descriptive framework, it matches the needs of a rule-based model very well.

The basic structure is the following:

A state in the task environment has to be identified from a set of observable cues, and a judgment made regarding the identity of this state. In the classical use of the model for experimental work within the social judgment paradigm, one of the problems is the match between the cues used for experimental work, and the cues actually used in the real life condition (as it is discussed in appendix 1 ). In the present application, this problem will be less pronounced, because the cues for judgment in a modern control room are related to discrete instrument indications. In the model, two sources of uncertainty are considered. One is cue reliability related to the fact that the "distal variable" which should be inferred from the cues may not be deterministically related to the cues available, a feature which may be used to represent, e.g., unreliable indications. Another is the cue utilization, i.e., the use of the available "proximal variables" for inference, reflecting the less-than-optimal use of cues by human judges.

The structure of the model proposed in the present approach is illustrated in fig. 3:

The relationship between the actual state of the process to be controlled and the cue set available to the operator, is perturbated (a) by the possible choice by the interface designer of a non-defining set of measured data, and (b) by the possibly unreliable measurement sensors. These perturbations represent the technical reliability of the control room indications.

Another perturbation function is used to represent the operator's cue utilization, i.e., to select the set of instrument or alarm messages which are included in the set by which the operator selects the task to perform. This perturbation function reflects the human error mechanisms or "judgment biases", and will be discussed in detail below.

The resulting, selected cue set is used to enter a decision table including the task repertoire, and a task is activated for execution. However, the task activated may not be successfully completed. Another perturbation function analyzes the match between elements of the current task and other tasks of the repertoire. If a match fulfilling certain conditions specified by the perturbation function is found, this task sequence will complete the (then erroneous) task sequence. The cues applied may not be quantitatively well defined, and fuzzy set models will probably be very well suited to represent cues, if the membership functions can be adequately formulated.

It will be seen that the model reflects the control requirements of the plant in terms of the proper task repertoire. Human features are represented only in terms of perturbation functions.

It will also be seen that the structure of the model matches the requirements of the DYLAM simulation system (Amendola et al. 1984, Cacciabue, 1986). The cue-set/proper-task data base represents a decision table model of the correctly performing operator and a specification of the perturbation function will

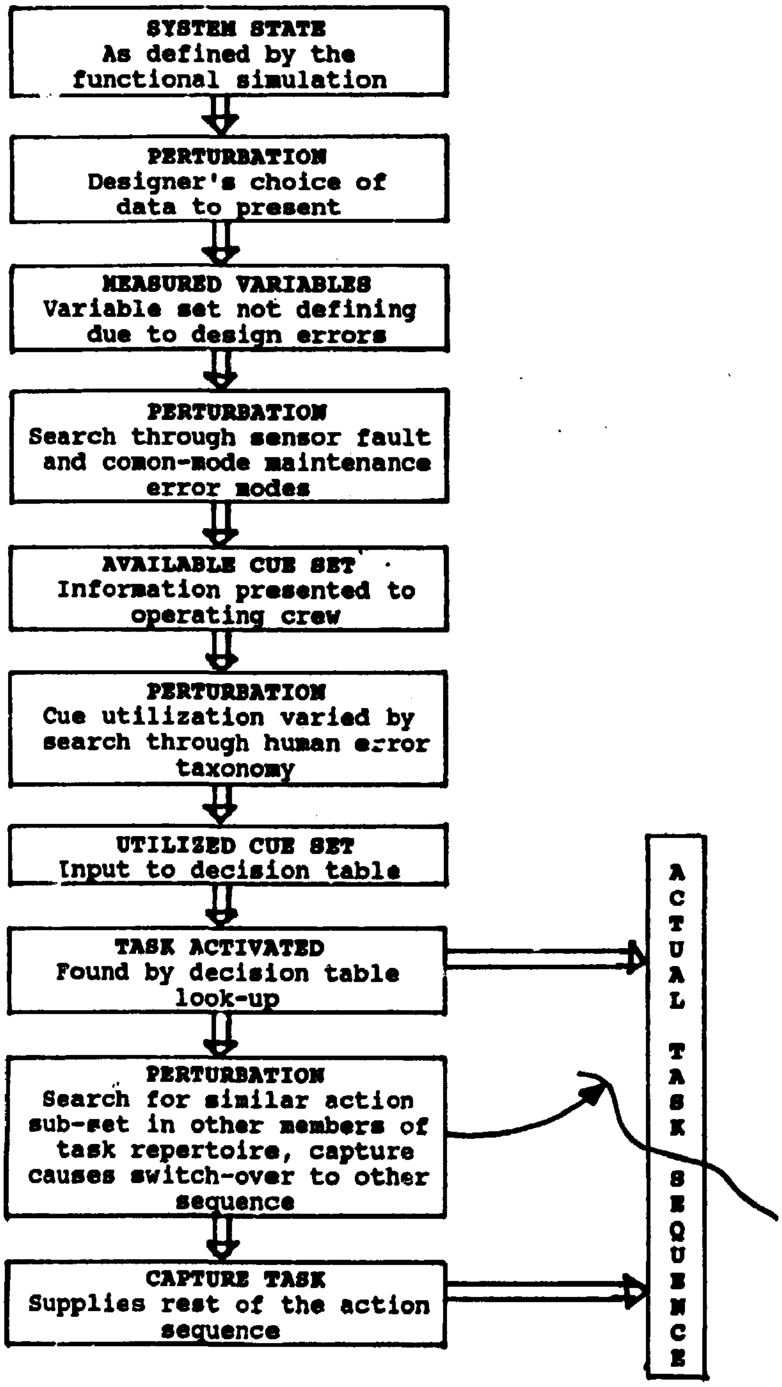



Figure 3. Information flow in simulation of operator responses in process control. Based on Brunswik's lens model and the human error taxonomy of Rasmussen et al., 1981.

make it possible in a systematic way to scan through the relevant error modes in an "error-mode-and-effect analysis similar to the one applied for technical component faults. The human element is represented by the perturbation functions which will be considered in more detail being a representation of the human error mechanisms.

In the previous sections it has been discussed that the model of the perfect operator is in fact a representation of the control requirements of the plant. The human element is "shining through" in terms of error mechanisms when adaptation breaks down. In the present modelling approach, the human element will be represented only in order to generate heuristic rules serving to limit the combinatorial search through the possible perturbations of the proper cue utilization.

The approach is based on the assumption that the systematic error mechanisms which are in focus according to the previous discussion, are related to interference between the actual task requirements during disturbed task condition and more frequently applied members of the task repertoire (Rasmussen 1985). This is in accordance with recent psychological research (Reason, 1985b). The review of error shaping factors, given by Reason et al. (1985) illustrates clearly the dependence on intertask or inter-situation interference.

The table reflects the relationship between different categories of human error, and modes of cognitive control. (See page 31).

Clearly, interference in a task repertoire will have a different basis in the different behavioral domains, a basis which can be related to different levels of abstraction in the behavioral control, and the skill- rule- knowledge- framework will be used to distinguish such levels of cognitive control (Rasmussen 1983). When we distinguish categories of human behavior according to basically different ways of representing the regularities of the behavior of a deterministic task environment, three typical levels of control emerge: skill-, rule-, and knowledge-based performance. These levels and a simplified illustration of their interrelation are shown in fig. 4. Skill-based behavior represents sensorimotor performance during acts or activities that, after a statement of an intention, take place without conscious control as smooth, automated, and highly integrated patterns of behavior. In most skilled sensorimotor tasks, the body acts as a multivariable, continuous control system synchronizing movements with the behavior of the environment, and performance is based on feed-forward control and depends upon a very flexible and efficient dynamic internal world model. Performance rolls along without any conscious choice between action alternatives.

| **BEHAVIORAL DOMAIN** | **ERROR SHAPING FACTORS** |
|---|---|
| SKILL-BASED | 1. Recency and frequency of previous use |
| | 2. Environmental control signals |
| | 3. Shared scheme properties |
| | 4. Concurrent plans |
| RULE-BASED | 1. Mind set ("It's- always-worked-before") |
| | 2. Availability ("First-come-best-preferred") |
| | 3. Matching bias ("like-relates-to-like") |
| | 4. Over-confidence ("I'm-sure-I'm-right") |
| | 5. Over-simplification (e.g., "halo-effect") |
| KNOWLEDGE-BASED | |
| | (Not considered for simulation at this stage) |

In general, human activities can be considered as a sequence of skilled subroutines composed for the actual occasion. The flexibility of skilled performance is due to the ability to compose from a large repertoire of automated subroutines the sets suited for specific purposes.

At the next level of rule-based behavior, the composition of such a sequence of subroutines in a familiar work situation is

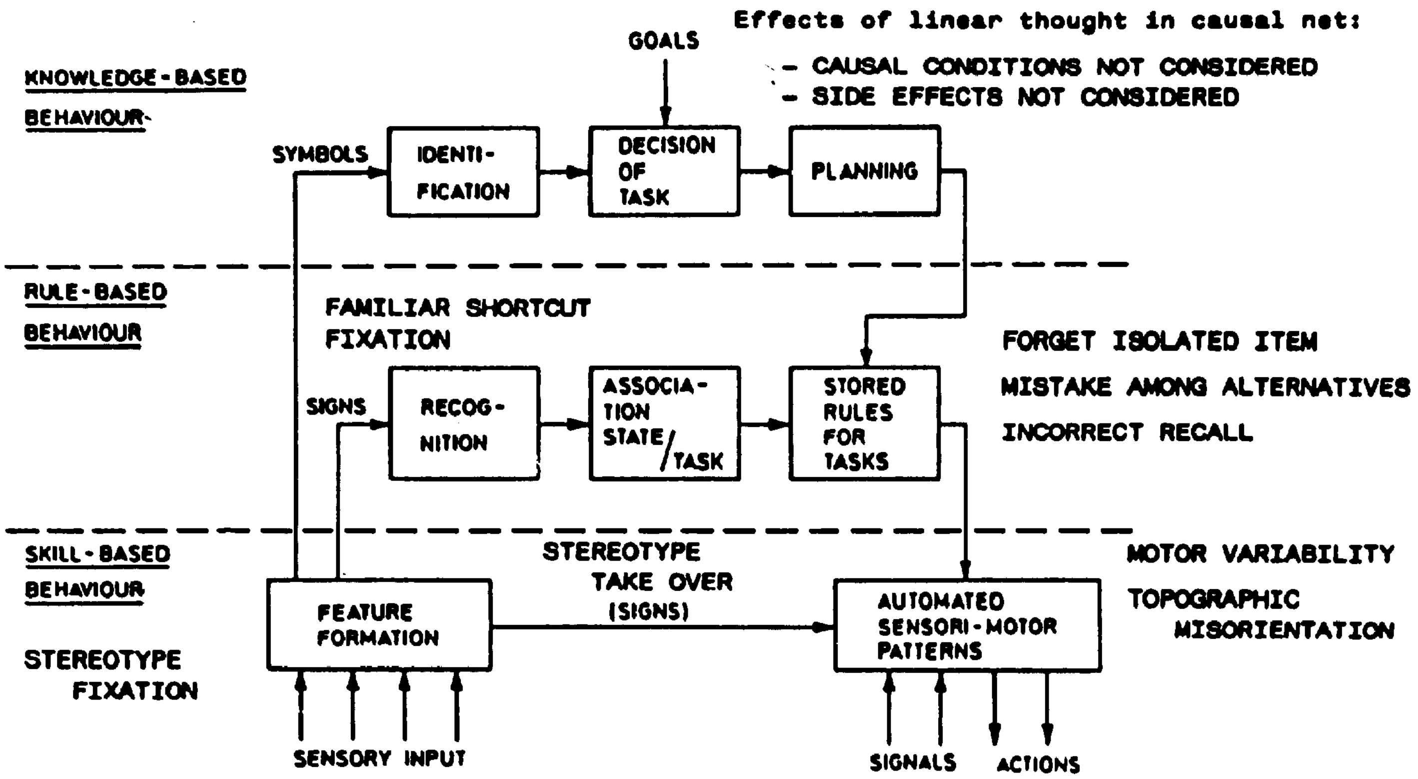



Figure 4. Typical human "error" mechanisms and their relation to control of behaviour. (Adopted from Rasmussen (1980) with permission from John Wiley & Sons, Ltd.).

typically consciously controlled by a **stored rule** or procedure that may have been derived empirically during previous occasions, communicated from other persons' know-how as an instruction or cookbook recipe, or it may be prepared on occasion by conscious problem solving and planning. The point here is that performance is goal-oriented, but structured by "feed-forward control" through a stored rule. Very often, the goal is not even explicitly formulated, but is found implicitly in the situation releasing the stored rules. The control is teleologic only in the sense that the rule or control is selected from previous successful experiences. The control evolves by "survival of the fittest" rule. In effect, the rule will reflect the functional properties that constrain the behavior of the environment, but usually those properties found empirically in the past. Furthermore, in actual life, the goal will only be reached after a long sequence of acts, and direct feedback correction considering the goal may not be possible. Feedback correction during performance will require functional understanding and analysis of the current response of the environment which may be considered an independent, concurrent activity at the next higher level (knowledge-based).

The rule-based coordination is in general based on explicit know-how, and the rules used can be reported by the person, although the cues releasing a rule may be difficult to describe. This level of cognitive control includes those situations when action alternatives are known, and a choice has to be made.

During unfamiliar situations, faced with an environment for which no know-how or rules for control are available from previous encounters, the control of performance must move to a higher conceptual level, in which performance is goal-controlled, and **knowledge-based** (knowledge is here taken in a rather restricted sense as possession of a conceptual, structural model). The level might more appropriately be called model-based. In this situation, the goal is explicitly formulated, based on an analysis of the environment and the overall aims of the person. Then a useful plan is developed - by selection, such that different plans are considered and their effect tested against the goal, physically by trial and error, or conceptually by means of understanding of the functional properties of the environment and prediction of the effects of the plan considered. At this level of functional reasoning, the internal structure of the system is explicitly represented by a "mental model" that may take several different forms.

It is clear that the basis for cognitive control is fundamentally different at these levels, and consequently so will be the mechanisms leading to task interference. At the level of skill-based performance controlled by sensorimotor patterns, interference in the topographical characteristics will be important. In the performance at the skill- and rule-based level, interference in the actual sign-patterns will be important (verisimilitudes), as well as interference between similar action seqences. Finally, in the knowledge-based domain, interference in the purpose/function/process/equipment mapping may be important, in addition to acts related to hypothesis testing during problem solving. Other kinds of error mechanisms may be relevant, but in the present it will context be particularly important to simulate the psychological mechanisms leading to systematic, causally activated errors.

It is important to note that the levels function as a hierarchical control system. Behaviour is composed of sequences of skill-based subroutines, which roll off as integrated smooth units without conscious control of the chaining of the individual acts when activated by the proper intention to act, <u>no choices</u> are to be made. At the rule-based level, skill subroutines are chained by <u>choosing</u> those suitable for the occasion according to know-how or prior experience. If a problem is at hand and no useful task sequence is available, knowledge-based "experiments" on a mental model may be necessary to <u>predict</u> outcome and <u>compare with relevant goals</u> in order to <u>generate a plan.</u>

The structure of this cognitive control necessitates the development of the simulation model "bottom-up", by first considering first the elementary skilled routines, and the basic error mechanisms of the skill-based control. Compared with Reason's GEMS approach, (Reason 1985) such skilled routines are activated as integrated task sequences by an intention or choice, and will be the elements of the task data base. More complex task sequences controlled at the rule-based level are composed of such routines, which are individually activated by a cue set including observation of indicators <u>and</u> the result of the antecedent act.

The basis of the simulation model will be the data base in the form of a decision table including the individual skilled routines, with the cue sets for activation as entries. The simulation of error mechanisms will be a perturbation function representing the systematic mechanisms behind errors. In the next paragraph, the perturbations necessary to simulate the error categories considered in the error taxonomy (Figure 5) will be discussed. Examples of operator errors illustrating the categories used are given in Appendix 2.

## MODELLING OPERATOR ERRORS

The perturbation functions to be included in the simulation model will be derived from consideration of the error mechanisms at the various cognitive levels. The model should be made by considering first correct performance of the frequent routine tasks, bottom-up in the skill-rule-knowledge model:

### Skill-Based Level

At the skill-based level, the operator is assumed to be completely adapted to the spatial and temporal characteristics of the interface. This means that instruments and keys are selected as intended. Then aberrations are introduced systematically in terms of search heuristics to identify opportunities for human

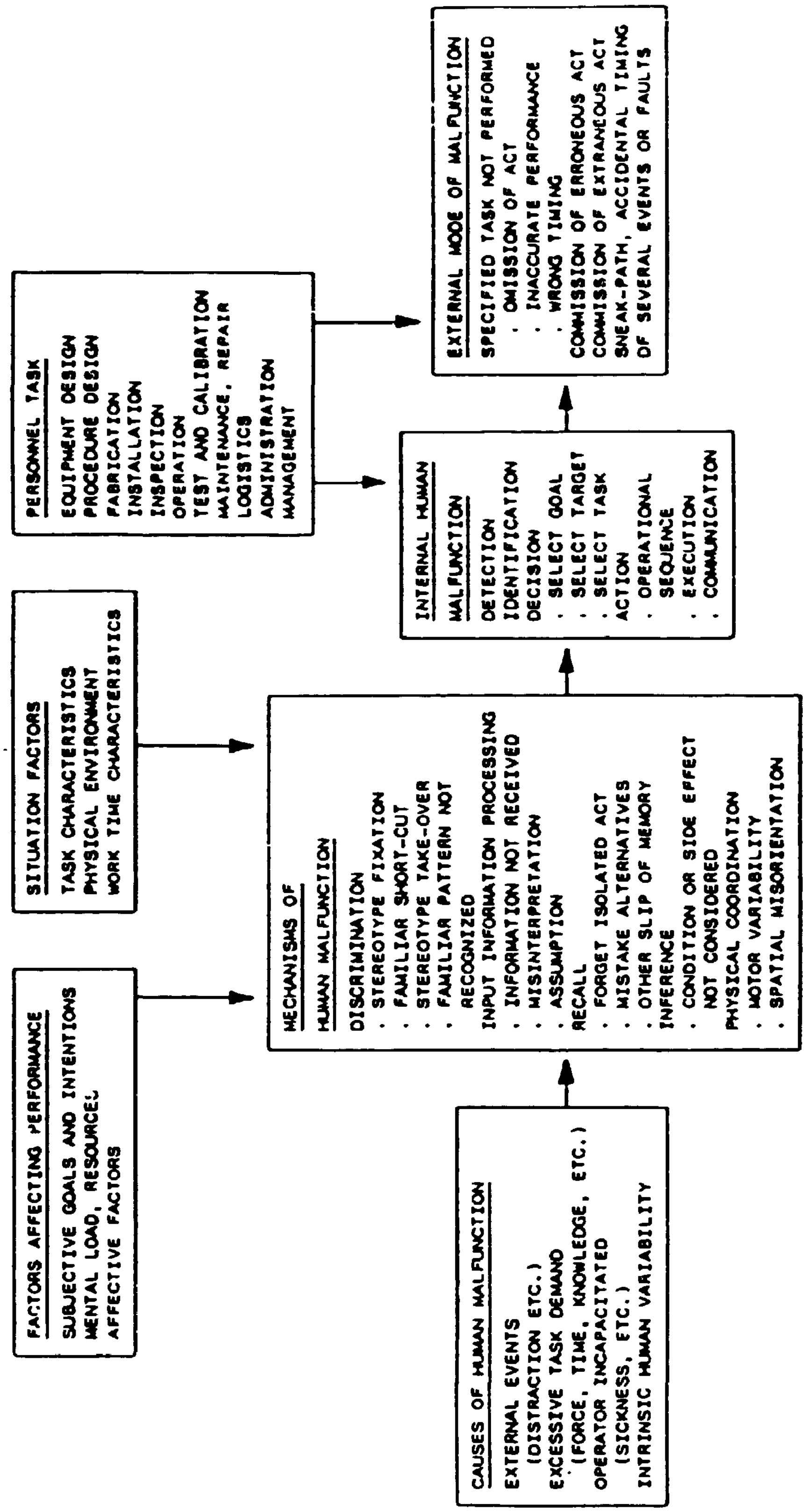



Figure 5. Multifacetted taxonomy for description and analysis of events involving human malfunction. (Reproduced from Rasmussen (1982) with permission from John Wiley & Sons, Ltd.).

errors (the classification labels of the taxonomy described by Rasmussen et al. (1981), are given in brackets). For each error mode, the heuristic rules to consider in order to delimit the need for combinatorially complete search during possible error modes will be indicated:

<u>Stereotype (skill) fixation.</u> (S1-1). This category represents the release of a familiar and normally very efficient work routine under inappropriate conditions. Ideally, the cue set for the task which is the simulated target for potential interference should be formulated and compared with the cue sets of all frequent activities, so as to identify potential for release of inappropiate tasks.

This is not possible, partly due to the high number of activities, partly due to the informal and holistic nature of the initiating cues at the skill-based level. Working backwards may be more realistic: The acts which may be sources of interference during a particular transient are identified by backward search (cause seeking) for sources of unacceptable acts on vital equipment. The frequent familiar tasks which include these acts as elements are identified by analysis of the action lists of procedures or interviews with skilled personnel, and the releasing cue set identified. Next, the cue sets of the accident scenario are scanned , to identify "close matches" which may release the unacceptable acts.

<u>Heuristic:</u> For the safety equipment in service during the transient, identify those work routines in the data base which involve these items or topographically close items. Match the related activating cue sets to the cue patterns present, and select those of near-match to activate as interference tasks.

<u>Stereotype take-over</u> (S1-3). Represents the situation when a task routine is properly activated, but another, typically more frequent, routine has overlapping subsequences, and takes over (capture error). Simulation involves a search for similarities between subsequences of tasks properly activated during the

transient, and other members of the task repertoire which may "capture control".

Heuristics: The task activated is compared with the other sequences in the task repertoire. Matching subroutine sequences are recorded and ranked according to size of overlap, and to frequency of occurrence. To limit the interfering tasks to consider, screening for lack of affinity of capturing tasks to safety equipment can be used (as above). Criteria for capture likelihood depending on frequencies and size of overlap should be found (preferably from operator metacognitive judgments?).

Motor variability (S5-1). This mode represents the lack of precision in manual tasks, and use of inappropriate force, etc. A probability distribution for the spatial precision as an overlay on the physical form of the interface will be able to generate errors of this category. The search will be extensive, unless it is limited to include only actions on keys which should not be operated, i.e., by looking only for highly familiar actions spatially close to critical items. Inappropriate force will not introduce new consequences (is covered by component fault).

Heuristics: For critical keys or items in the safety equipment in operation, scan the task data base to identify operation on items in the topographic vicinity. Likelihood is graded by distance.

Inadequate spatial orientation (S5-2) represents the cases when operators lose their orientation while moving around, and operate on equipment at a wrong location. This depends on similarities in spatial arrangements and appearance and will not be covered by the simulation because the underlying configurative information will not be available. The effects will be partially covered by the "mistake of alternatives" at the rule-based level.

**Rule-Based Level**

At the rule-based level, the normal, successful performance will be in terms of a "production system" which represents the procedural rules activated by cues derived from the defining attributes. The model will be a logical complete version of the Brunswickian lens, i.e., "naturalist guide", implemented in a decision table, which reflects only the control requirements of the system. Actual human performance will depend on convenient signs, rather than defining cues.

An important consideration in the modelling effort will be to investigate the possibility of systematically generating the likely set of convenient sign developed by skilled operators. The basis for such sign identification will be an application of some basic psychological principles controlling human adaptation.

Several studies (e.g., Bruner, 1969; Rasmussen et al., 1974) indicate that skilled individuals adapt to strategies reflecting preference for "the way of least resistance", minimizing cognitive strain and load on short-term memory. From such general principles, very specific tactical rules can be derived, which will be efficient guides in the heuristic delimitation of the necessary variations of cue perturbations to include in the simulation model. Another basic principle will be that human memory is based on generative mechanisms, not rote memory and, therefore, items not integrated in a meaningful whole are likely to be forgotten.

In the following paragraphs, heuristics for generating the perturbation functions will be discussed with reference to the rule-based error modes.

**Familiar association short-cut** (S1-2). This error mode represents the tendency of humans to identify very effectively correlation between the requirement for a certain act and a small set of convenient cues which may be completely informal (relay and motor noise, certain easily read instruments, etc.) and not intended for that purpose. In a modern control room, cues are most likely to be associated with instrument readings, and, therefore, to a first approximation the "sign" adopted can be expected to be a sub set of the instrument indications which form the designed state defining set. When this subset is no longer valid as an activating cue, because the system condition has changed, a wrong task is activated. If, for instance, the cues normally activating a routine task are a subset of the cues related to an emergency procedure, there is a risk that the routine task will be inappropriately performed during an emergency, in particular during early phases when an emergency is not yet perceived by th operators.

The perturbation function used to represent this error mode will have to scan the available cue set systematically during the situation considered in search for a match with the cue (sub)set related to the task repertoire. In order to avoid the combinatorial explosion in this search, an effective set of heuristics representing the operators' tactics in reading instruments should be applied. It should, however, be remembered that it is not a reliable model of human behavior that is needed, but a model describing a well defined envelope within which the systematic mistakes which are designed, trained, or instructed into the system will be found.

Different heuristics for limiting the search can be proposed from more basic adaptive properties:

Humans are economic and seek the way of least effort. Therefore, it can be expected that no more cues will be used than are necessary for discrimination among the tasks belonging to certain scenarios or situations which can be expected to guide "process feel", i.e., operator expectations. A crude representation can probably be made by defining "operation modes", or situations - analogous to Minsky's frames - which will be labels for a related set of tasks that are then more likely activated than tasks outside the set. The formation of economic, convenient cue sets would then only have to be discriminative within the particular set. When situations change, reliance on the cue sub-set which is no longer valid, will cause interference due to inappropriate "expectations".

The concept of "process feel" or internal model is compatible with the basis of the lens model (compare with arguments of Tolman and Brunswick mentioned above): Operators are not accepting "input" or responding to stimuli, they are actively asking the question to the system which will be necessary to direct or confirm their expectations.

This implies a systematic way of identifying the cue subset which operators will adopt. Sets of tasks which are related to the same work scenarios or plant operating modes are selected from the task repertoire, and the cue sets which are necessary and sufficient for task discrimination within the scenario are identified. These sets will represent the convenient signs related to the tasks. The same task will be related to different scenarios, and, therefore, several cue sets may be related to the task, probably with a primary joint set and some additional qualifiers. An analysis of the total cue set and the assumption of mental economy will lead to hypotheses about the most likely scanning sequence of the cue set for a given scenario, such as scanning according to the most salient cues of tasks considered in descending order of frequency of occurrence, or scanning in the order of most discriminative effect.

The proper hypothesis to choose (perhaps several scanning principles should be included to get a reliable "envelope") can be decided by interrogating operators, since the information needed seems to be accessible by Reason's metacognitive judgment techniques. The work scenarios or operating modes to consider should be identified from work schedules, operation logs, and operator interviews. Comments to the present approach from social judgment research which is the origin of the lens model (Brehmer, 1985, private communication) have implied that it should be considered, when comparing with results from the classical experimental use of the lens model, that the focus of those subjects is normally given by the experimental condition. This is not the case in control room emergency situations. However, the definition of operating modes, and the effect of the operator's process feel may add up to similar conditions.

Another guide for the scanning strategy can be derived from the striving for pattern for matching and the principle of "the point of no return". When information is observed sequentially, evidence (Rasmussen et al., 1974) seems to show a strong tendency to make a decision to act as soon as information is pointing to a familiar routine. Even when subsequent observations indicate that the routine will be inefficient, the decision will not be reconsidered. This points to a search for interfering tasks by matching cue subsets of the emergency situation to the task repertoire in the <u>order</u> of cue observation. A task is then activated from a match between the most salient and frequently used cues, irrespective of mismatch in the remaining part of the set.

<u>Heuristics:</u> The empirical rules applied to avoid the need for exhaustive search will be related to choice of the most likely scanning order of cue observation, and a match of cue subsets to subsets of the entry vector of the task data base. The activating cue for the task data base will be determined from an analysis of the set necessary for discrimination within "operating modes." Interrogation of operating staffs will be useful.

<u>Information not received/sought</u> (S2.1). This category can be identified by a search similar to the previous one. Cues are systematically omitted from the set, and the match to an activity by the rest-set is evaluated. The category will probably include omission of observation of "normally redundant" cues, and will therefore, to a large degree, be covered by the category just mentioned.

<u>Misinterpretation of information</u> (S2.2). Members of this category may be very difficult to identify, since misinterpretation will depend on subjective expectations. Some members may be found by

replacing cues by other cues obtained by "one-element-substitution" , e.g., moving decimal points, replacing letters, etc.

Heuristics: A useful rule for limiting the search to perform may be to look for similarities between a cue pattern related to an unfamiliar situation, and compare with the cue patterns of tasks which are normally performed in the particular operating mode and which may activate an expectation leading to misinterpretations. In a way this heuristic rule adds to the "familiar shortcut" mode, an erroneous match which is independent of the scanning order.

Forgetting isolated acts (S3.1). All action sequences can be analyzed to identify isolated acts which are likely to be forgotten. In particular, it should be done when the same action sequence is used repeatedly, and an omission can lead to coupled errors. This error mode can only be included from an analysis of the content of the particular tasks that are included as targets for interference which, however, will be a limited number. The analysis will require technical insight, since it will depend on knowledge about the actual verbal labelling of the task (cf. Worf: "the name of the task affects behavior") and about the effect of omissions on the rest of the sequence.

Mistake among alternatives (S3.2). All tasks are analyzed to see whether the cue or action sets have members for which immediate alternatives are available, such as numerical signs, numbers, directions, etc. A search procedure is then designed to scan this list when the task is activated.

Other slips of memory (S3.3). Can be anything. Most effects are probably covered by categories of technical faults in the fault tree analysis, since systematic coupling to other events is unlikely (should be identified by the other categories, or they should be identified by " backward identification").

**Knowledge-Based Level**

At the knowledge-based level where control of behavior depends on problem solving based on operations on a mental model of the system to control, the variety of error modes which have to be identified and formulated for inclusion in the automatic scanning during simulation, will be much larger than in the skill- and rule-based levels. This is because the cognitive processes involved will be much more varied. In addition, the mental model available to the operators will not be formed and maintained in a uniform way by the daily work requirements, and therefore cannot be found by an analysis of the system requirements as it is the case for the task repertoire influencing the error modes at the rule-based level.

Automatic control of a failure-mode-and-effect analysis at this level is at present considered less realistic. However, some categories of errors may be identified by searching for interference between mental models at different levels in the purpose/function/process/equipment hierarchy, in terms of competing means-ends relations of the same person or competing intentions of different persons.

At present, it is considered an advantage to test the simulation model at the skill- and rule-based level before the model is complicated further by including functions at the knowledge-based level. If a situation appears during simulation when such performance is required because no match, or near-match, is found between the cue set present and the entry in the task repertoire, an interrupt to the experimenter may be made,as a first approximation and, in a way, simulating an operator requiring the advice of a supervisor.

**Comparison with Other Taxonomies**

During recent years, other taxonomies of human error and bias of judgment, based on the consideration of cognitive control mechanisms have been proposed, and a comparison between the error categories will be useful. In this section, a mapping is considered of the human error and bias categories suggested by Reason and Embrey (1985) onto the simulation model:

**Error shaping factors at the skill-based level**(p.78)

Recency and frequency of prior employment. "The more recently and frequently a particular routine is set into motion and achieves its desired outcome, the more likely it is to recur uninvited as a slip." This mode is equivalent to the stereotype (skill) fixation (S1-1).

Environmental control signals. "Familiar environments trigger associated action routines, particularly at moments of reduced intentionality. These erroneous actions obey local rules even though they are out of step with the current plan." This error mode is closely equivalent to the stereotype takeover mode(S1-3).

Shared scheme features. "From the pattern of slips observed in everyday life, it is evident that the intentional activation of a given set of action schemata also has the effect of increasing the activation of other schemata possessing shared features." This is another way of presenting the stereotype-take-over failure mode (S1-3).

Concurrent plans. "Many slips betray the influence of concurrent plans or intentions. These can take the form of blends in which two active plans become intermingled in the same action sequence." To a first approximation this mode is also taken care of by the stereotype takeover mode of error (S1-3). (In the present context, the simulation of the error mechanisms which are likely to give high likelihood couplings is the prime consideration.)

**Error shaping factors at the rule based level** (p.80):

Mind set (Einstellung). The fixation mechanisms described by Luchins and Luchins (1950), who experimentally demonstrated the potential rigidity of rule-bound behavior. This error mode is directly equivalent to the "familiar short-cut" (S1-2).

Availability The error inducing effects of the availability bias have been demonstrated by Tversky and Kahnemann (1982). "Things that come more readily to mind are likely to be more frequent, more probable, more important, more useful, and better understood than less readily available items." This heuristic bias may be considered by putting recency or frequency weights the search function, together with the selection function related to cue-subset matching. A fuzzy rule like "recently-used-in-this-context" weight. Another fuzzy consideration may be "this task is normally following now", related to process feel.

Matching or representativeness Bias. This bias is related to the fact that human judgment is based on prototypical features, rather than defining attributes, and is caused by matching a prominent feature of a situation to a stored situational representation also possessing that feature. This matching bias will also be well represented by the "familiar short-cut" mode.

Over-simplification. "Studies in which people are asked to rate objects or other people on each of several specific dimensions reveal a striking inability to take account of the independent way in which these entities vary along the different dimensions. Instead, there seems to be an overwhelming tendency to collapse all these dimensions into a single ordering by merits or 'good versus bad' dimension thus losing valuable information about the pattern of attributes unique to a single object or individual." There seems to be no directly equivalent to this error mode in our taxonomy. However, the behavior tendency implied is the same as we have found in our highly skilled routines, and whether they constitute errors or a specific, effective (focussing) strategy, depends on the context. It may be argued that the heuristics is in particular used in knowledge-based control.

Over-confidence. "Problem solvers have a marked tendency to be over-confident in evaluating the correctness of both their rule selections and their knowledge of the system." This formulation in fact indicates a relation to the knowledge-based domain, if not it is a manifestation of the 'familiar short cut' routine, i.e., confidence in inductive generalization.

In conclusion, a structured mapping between the error categories of the various cognitively based taxonomies can be formulated and, therefore, the taxonomy proposed by Rasmussen et al. (1981) can be considered a suitable starting point for the development of a simulation model.

**Error Recovery**

Error recovery is a very important feature if probabilities of a particular scenario are to be estimated. Error recovery depends on some kind of error detection followed by corrective actions. Detection may be due to "the wrong key effect", i.e., the continuation of the intended task sequence is impossible, unless the error is corrected. This depends solely on features of the plant. Detection may be due to an operator realizing that a task has an effect in contrast to his expectations. This requires modelling of his expectations as is included in optimal control models by means of a Kalman filter representation of plant performance. It is, however, a question whether operator expectations about the response of dynamical, technical systems depend on a dynamical internal model of the entire plant to be controlled. This may be the case when operators are in a tight manual control mode, but in case of occasional control during disturbances, the "expectations" may depend on a more fragmentary internal model, composed by representations of dynamic entities at a higher functional abstraction level in terms of categories of prototypical "time-constants" of typical elements such as thermal systems and rotating systems , and only judged by, for instance, their physical sizes.

Furthermore, it is a question to what extent the human detection function need to be represented. The "expectations" in fact reflect system properties, the time instant when detection can occur is determined by the dynamical time delay in the system; whether it will occur depends on the preoccupation of the operator, which means that both branches (recovered and not recovered) should be represented in the risk analysis. The quality of an operator model will only affect probabilities.

The "time-constant" itself of the system operated on will determine the time interval to the most likely error recovery by a skilled operator, and thus the time for considering branching in the event tree. An important aspect to consider at this point in time may be the possibility of another erroneous operator act at the time of error detection, if the cue pattern present matches (or nearly matches) an unacceptable action sequence. Again it should be remembered that only the systematical error modes causing extra couplings in system structure are to be considered. Therefore, features peculiar to a particular, individual performance related to unsystematic error recovery should not be included.

## THE RESULTING MODEL

In summary, the structure of the model resulting from these considerations will be the following:

The state of the system will be represented by a set of measured variables. From this set, a set of cues for operator actions is generated by means of a perturbation function representing the cue validity in the Brunswick lens or sensor reliability in the technical sense. During simulation, a control program will scan through the relevant perturbations.

The resulting cue set represents the observations available to the operator, and will be transformed by another perturbation function representing the Brunswickian cue utilization. A control program will have to scan through all relevant perturbations by heuristics derived from the error taxonomi discussed above.

The cue set thus obtained will be the entry pattern for a decision table search for a matching task sequence, which will be retrieved from the task data base and, in turn, define the action sequence to be fed to the plant simulator.

In addition, a search for match between action subsequences of the task retrieved and subsequences of other, typically more frequent, tasks should be performed. In case of a match, a switch-over to the similar task is simulated after the common subset.

The structure of the model is very similar to the basic component model of DYLAM: The operator, as a component, is represented by a decision table associating cue sets to proper action sequences. Search through the relevant error modes is controlled by a set of perturbation functions controlled by heuristics derived from the available error taxonomies. In the DYLAM context, an attempt is made to represent the operator responses in terms of a fuzzy set model representing errors in terms of verisimilitude of membership functions of cues (Mancini, 1986). The mapping of the various error categories discussed in the previous section in terms of such membership functions will be useful for system development.

<u>In conclusion, a stepwise development of a model of operator responses to disturbances of plant operation based on the modified Brunswickian lens model appears to be a feasible approach. A program for testing the model should be planned.</u> One approach for testing will be to use the well structured micro-worlds of computer games, which offer a context in which the entire task repertoire can be formulated, and a consistent data collection as well as data analysis can be made. (Rasmussen, 1985). Another possible test bed would be the GNP (Generic

Nuclear Power) experimental facility (see Goodstein 1983). In these experiments, the model should simulate a computer game player and an operator of a simulated GNP plant, and the performance (in particular the error modes) compared with recordings of human players, operators respectively.

Finally, as a test could be used would be a comparison between the event trees that would be identified by the simulation model and those found by human analysts in a benchmark test of procedures for industrial risk analysis.

## CONCLUSION AND RECOMMENDATIONS

The main conclusion of the review presented in the present report is that a plan can be formulated for a systematic development of a model suited for simulation of operators' responses to disturbed process plant operation. A prerequisite for such a plan to be realistic will be a coordinated set of assumptions which take into consideration the special features of the design practice for potentially risky process plants, of the procedures for probabilistic risk analysis, and of the basic psychological mechanisms bringing structure to the error modes of skilled operators.

Development of the model will depend on several lines of research dealing with separate topics. These should, however, be carefully coordinated to reach a useful result with a reasonable demand on research resources:

A careful consideration of the requirements of the risk analysis which is to be supported by the model can identify feasible, simplifying assumptions with respect to the situations and tasks for which responses should be included. Several stages of the development can be envisaged. In the first approximation, the model will serve only to identify responses adding new branches to the fault trees which have been identified from considering merely the technical component faults. In addition, simplifications can be made by looking for models creating an envelope known to include the relevant operator errors. The resulting error modes should then be screened for irrelevant modes by the risk analyst. The basic philosophy is that a well formulated boundary of the model will be more important than precise coverage. The model should then be refined gradually by more precise definition of the limits of coverage, together with consideration also of tasks which only contribute to the probability of already identified events. These aspects imply an analysis of the detailed requirements of the risk analysis that is to be supported by the model, and of the various possible, simplifying assumptions.

Another major line of development will be the **development of a systematic model of the task repertoire** of the crew operating the actual process plant. The aim is the development of a tool which will serve to identify the interference between the different tasks which can lead to systematic coupling between otherwise independent events, i.e., common mode errors which are in a way unintentionally designed, trained, or instructed into the system. From this point of view, a systematic data base including characteristics of the total task repertoire appears to be as important for a risk analysis as is a data base including component characteristics and fault nodes. Three lines of development should be considered and tried out for this purpose:

1. An extension of the existing task analyses, in particular with respect to identification of the observation cues which are used by the crew for control of the tasks.

2. Considering that a model of a properly performed task is in fact a model of the control requirements of the process plant, the fundamental structure of a given task can be derived systematically from a representatiom of the functional properties of the plant. A program should be formulated to test how far this approach can serve the

development of a task model. A prerequisite will be a consitent representation of the functional properties of the plant in terms, for instance, of a means-ends hierarchy including causal relations as well as information on design intentions. This implies a separate data base development which, however, has also been identified as a basic need for the design of decision support systems in the plant control rooms.

3. Finally, methods should be developed and tested, for the interrogation of operators, in order to get information about the actually used task sequences and, in particular, about the observation cues, the cue scanning tactics, and the frequency of the different tasks. For this purpose, methods for "expert knowledge acquisition" considered for design of expert systems, and methods for "metacognitive judgement" should be tested.

In the first phase the aim will be to develop a model including the skill- and rule-based task performance. In a later phase, also **the knowledge-based domain should be considered**. The review of cognitive models available has identified several model approaches which can be the basis for this development. They will all need a representation of the operator's "mental model" in one form or the other and, to prepare for this extension, a program for the **study of mental representations of physical systems** will be needed, theoretically as well as experimentally. Such a study can advantageously be coordinated with the formulation of the means-ends hierarchy mentioned above which can serve as a framework for identification of the content (not the form) of the mental model. Related to this work will be the study of the benefit to be gained from **use of fuzzy set and optimal control models** for representation of mental processes and "process feel", respectively. This extension will be important in particular for a representation of error recovery features which will be necessary for a more refined model (i.e., a model with a narrower envelope around the relevant performance sequences), and which will open up for the practical development of the integrated models described in the review.

Finally, an important part of an integrated program will be a plan to **test the model concepts**. Different stages in such a test program can be proposed:

1. Since the quality of the model depends entirely upon a representation of the entire task repertoire, it will be important to test the model concepts related to the error mechanisms separately in a task context for which a well bounded set of task sequences can be identified. For this purpose, a well chosen computer game appears to be suited, in particular because schemes for systematic data collection and analysis can be developed. For this purpose, the cognitive task characteristics of various games should be compared with the requirements of the process plant tasks, and an experimental program as sketched in Appendix 3 considered for the first test.

2. Next, a test of the validity of the task models developed by the approaches proposed above should be performed by a simulation of a simplified process system (as, for instance, the GNP system) in a laboratory environment where systematic methods for observation and data analysis can be used by trained experimental psychologists.

3. Finally, when an acceptable task model is available, a full scale test should be made by a simulation together with the DYLAM system as a part of a bench mark test in which the simulated performance is compared with the event sequences as identified by a panel of expert risk analysts. From the structure of the model and of DYLAM, no conceptual problems are envisaged; however, only a set of algorithms will be needed for control of the perturbation functions of the present model in synchronism with the DYLAM.

In all, a promising line of research can be entered, combining recent results from cognitive psychology, risk analysis, control theory, and process plant simulation. A successful program will, however, require a coordinated approach to all the elements of the problem concurrently by an interdisciplinary team.

**REFERENCES**

Amendola A., (1984). DYLAM-1 A Software Package for Event Segence and Consequence Spectrum Methodology. EUR 9224 EN. Ispra: EEC-Joint Research Center.

Burgy, D. ,C.Lempges, A.Miller, L. Schroeder, and H. vanCott, (1983). Task Analysis of Nuclear Power Plant Control Room Crews, Vol.1-4. NUREG/CR-3371, Washington DC: U.S.Nuclear Regulatory Commission.

Bruner, J. S., J. J. Goodnow, and G. A. Austin. (1956). A Study of Thinking. New York: John Wiley & Sons.

Brunswick, E. (1957). Scope and aspect of the cognitive problem, in: Contemporary Approaches to Cognition. H.E.Gruber et al. (Eds.). Cambridge: Harvard University Press.

Cacciabue, P. C. (1986): System Response Analyser: A Tool for the Analysis of the Behaviour of Systems. Tech. Note No. 1.05.C1.86.12 PER 1100/86. Ispra: Joint Research Center.

Goodstein, L. P., J. Hedegaard, K. Søe Højberg and M. Lind. (1983). The GNP Testbed for Operator Support Evaluation. Roskilde, Denmark: Risø National Laboratory, Report No. M-2460.

Hammond, K. R. (1966). The Psychology of Brunswik. New York: Holt, Rinehart and Winston.

Lind, M. (1981). The use of flow models for automated plant diagnosis, in Human Detection and Diagnosis of System Failures, J. Rasmussen and W. B. Rouse, eds. New York: Plenum Press, 422-432.

Lind, M. (1982). Multilevel Flow Modeling of Process Plant for Diagnosis and Control. International Meeting on Thermal Nuclear Reactor Safety, Chicago. (Also: Roskilde, Denmark: Ris National Laboratory, Report No. M-2357).

Mancini, G. (1986): Modelling Humans and Machines. In: Hollnagel, E., Mancini, G., and Woods, D. (Eds.): Intelligent Decision Aids in Process Environments. Berlin: Springer Verlag, in press.

Miller, R.B. (1963). Task Description and Analysis. In: R. M. Gagne (Ed.): Psychological Principles in System Development, 187-228. New York: Holt, Rinehart and Winston.

Pedersen, O.M. (1985). Human Risk Contributions in Process Industry: Guides for Their Pre-Identification in Well-Structured Activities and for Post-Incident Analysis. Risø-M-2513, Roskilde: Risø National Laboratory.

Rasmussen, J. (1980). What can be learned from human error reports, in Changes in Working Life. K. Duncan, M. Gruneberg, and D. Wallis, eds. London: John Wiley & Sons.

Rasmussen, J. (1982). Human factors in high risk technology. In: E.A.Green (Ed.). High Risk Safety Technology. Lodon: Wiley and Sons.

Rasmussen, J. (1983). Skills, Rules, Knowledge; Signals, Signs, and Symbols; and Other Distinctions in Human Performance Models. IEEE Transactions on Systems, Man, and Cybernetics, Vol. SMC-13, No. 3.

Rasmussen, J. (1984). Human Error Data. Facts or Fiction? 4th Nordic Accident Seminar. Rovaniemi, Finland. Risø-M-2499.

Rasmussen, J. (1985): Use of Computer Games to Test Experimental Techniques and Cognitive Models, Third Symposium on Empirical Foundation of Information and Software Science, Risø, Roskilde, October 1985. To be Published.

Rasmussen, J. (1986). On Information Processing and Human-Machine Interaction: An Approach to Cognitive Engineering. New York: Elsevier Scientific Publishers.

Rasmussen, J., and A. Jensen. (1974). Mental procedures in real life tasks: A case study of electronic trouble shooting. Ergonomics. 17, (3), 293-307.

Rasmussen, J., O. M. Pedersen, A. Mancini, A. Carnino, M. Griffon, and P. Gagnolet. (1981). Classification System for Reporting Events Involving Human Malfunction. Roskilde, Denmark: Risø National Laboratory, Report No. M-2240.

Rasmussen,J., Pedersen O.M. (1984): Human Factors in Probabilistic Risk Analysis and in Risk Management. In: Operational Safety of Nuclear Power Plants, Vol.1, I.A.E.A., Vienna.

Reason, J. (1982).Absent-Minded? Prentice- Hall.

Reason, J. (1985a). General Error Modelling System (GEMS). In: Rasmussen, Duncan and Leplat (Eds.): New Technology and Human Error. Proceedings of 1st Workshop on New Technology and Work. Bad Homburg. John Wiley. To be published.

Reason, J. (1985b). Simillarity-Matching, Frequency-Gambling, Inference, and Metaknowledge: Their Relative Contributions in Different Memory Search Tasks. (Working title). In preparation.

Reason, J.T. and Embrey, D.E. (1985). Human Factors Principle Relevant to the Modelling of Human Errors in Abnormal Conditions of Nuclear and Major Hazards Installations. Technical Report. Dalton, UK: Human Reliability Associates Ltd.

Reina, G. and A.Amendola, (1981). DYLAM: A Computer Code for Event Sequences and Consequence Spectrum Analysis. Int. ANS/ENS Top. Meeting on Probabilistic Risk Assessment, Sept. 20-24, 1981, Port Chester N.Y.

Tolman, E. and E Brunswik (1935). The Causal Structure of The Environment. Psychol. Rev., 42, 43-77.

**APPENDIX 1**

# COGNITIVE MODELS OF HUMAN BEHAVIOUR

## A LITERATURE REVIEW

Jens Rasmussen

Edited extract from Jens Rasmussen: On Information Processing and Human Machine Interaction: An Approach to Cognitive Engineering, Elsevier Scientific Publishers (in press)

APPENDIX 1

**REVIEW OF COGNITIVE MODELS**

The skill-, rule-, and knowledge-based framework is useful to structure a description and discussion of various domains of human behavior and to distinguish among them. Because the control mechanisms of the different categories, as well as the interpretation of information from the environment, are basically different, it is likely that different types of models are suitable for quantitative prediction in the various domains. Analytical models can be very important in the design of human-machine systems for optimizing task performance, in particular, for rare event scenarios which are to be considered for design in centralized, high risk installations and for which system performance cannot be verified experimentally. A problem one has to face in such situations is that modeling of the interaction of categories of behavior is very important, and this requires compatibility among the separate kinds of models. It is not intended here to go into detail on the various models, but instead to review their main characteristics in order to select a model for the present specific purpose.

The natural starting point of this discussion is a review of the approaches toward the modeling of the subconscious, skilled interaction with the physical environment. Behaviour in this domain has some characteristics that determine the nature of useful models. Behaviour is not controlled by a set of process rules, but by a dynamic, internal model of the environment which controls the sensorimotor interaction by means of signal processing in the space-time domain. Because skilled performance only includes behavior of well-adapted persons, models of

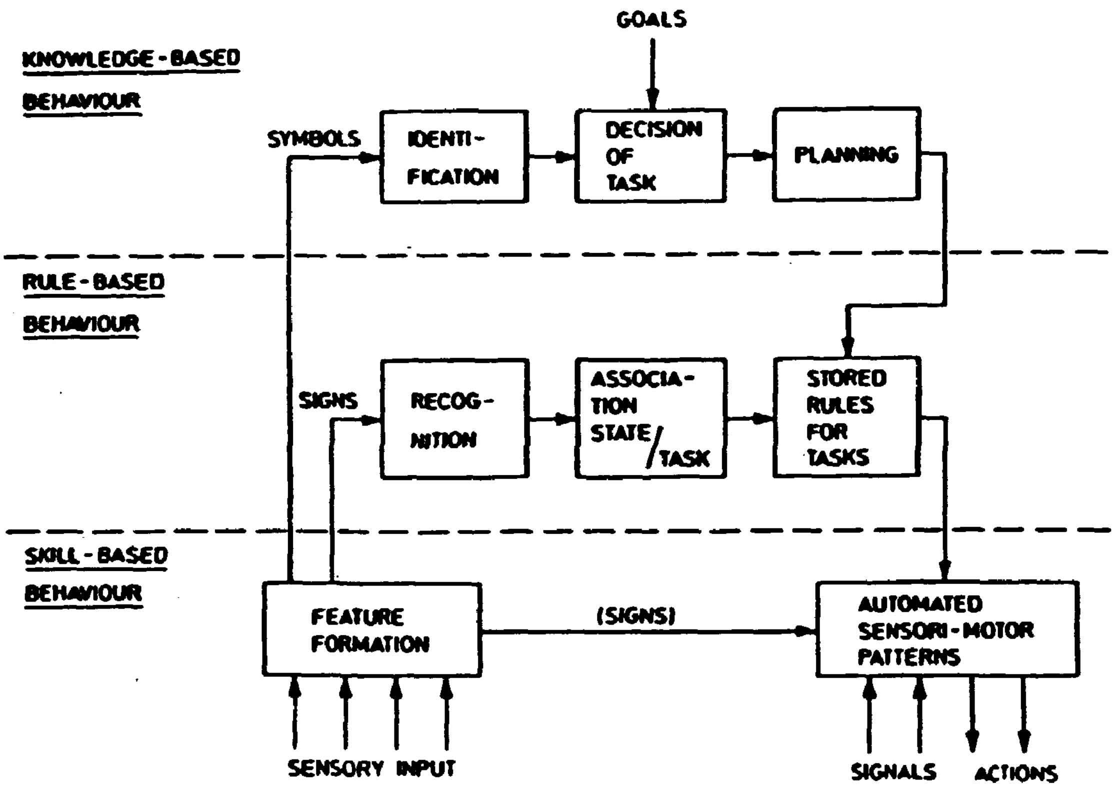



Figure 1. Simplified map of three levels of control of human actions. The three levels are in a real situation interacting in a much more complex way than shown. Note that the arrows indicate information flow, not necessarily control. For instance, only the lowest level can be considered "data-driven". For the higher levels, information is actively sampled ("rule"- or "model-driven"). However, two aspects are necessary in explaining human behaviour: Reasons in terms of intentions, conditioning the organism top-down, and causes in terms of sensory information releasing actions bottom-up in the figure.

. (Reproduced from Rasmussen (1983) with permission from I.E.E.E.).

performance in this domain will generally be models of features of the environment as viewed through human selectivity and limitations. If models should also include the quantitative precision necessary in control of manual acts, sets of mathematical time functions seem to be the only reasonable choice. Models in this category have been discussed in detail by Sheridan and Ferrell (1974). Two aspects of skilled behavior are to be considered, and have typically been considered separately, in modeling. One is the control of human attention allocation, the other aspect is the control of the manual interaction.

## Attention Allocation

Humans do not constantly scan the environment and extract meaningful features from the available flux of information. Acting in a familiar environment, people sample the environment, controlled by their expectations as to where an update of their internal dynamic world model is needed. This means that this model specifies when an update is needed, and where to look. Different approaches to model this function have been tested experimentally.

One family of models is based on queueing theory. The system considered in queueing theory is a person serving a number of tasks. The tasks cannot be attended simultaneously, but have to be considered on a time-sharing basis according to a service strategy depending on the nature of the tasks. Many task demands, such as instrument reading during a monitoring task, arrive randomly. Typically, queueing theory considers demands with poison or exponential distributions. Queueing models of attention allocation postulate that humans optimize their performance according to a service strategy considering the arrival sequence and task priority. Queueing theoretic models have been used by Carbonell (1966) and Carbonell et al. (1968) for a study of instrument scanning behavior of aircraft pilots in order to predict the fraction of time devoted to each instrument. Also, Senders and Posner (1976) have developed a queueing model for monitoring tasks. Rouse (1977) has employed queueing theory to model pilot decision making in a multitask flight management task. Queueing models basically represent the time distribution and priority characteristics of the task environment and can therefore be useful for analysis of the workload posed in terms of time and scanning requirements in a monitoring task.

Another approach in the frequency domain is based on Nyquist's information **sampling theorem** which states that the information from a source having spectral components with an upper limit frequency of **w** Hertz can be completely represented by an observer who samples **2w** times per second. The sampling model has been tested by Senders (1964) in experiments where the subjects' task was to respond to a number of instruments fed by random signals of different bandwidth. Also data from pilots in real flight tasks seem to match the model (Senders, 1966). These findings support the statement that highly trained operators in fact mirror the properties of the environment as long as performance is well within capacity limits.

## Detection

The ability of humans to detect changes in the behavior of a system that they are monitoring is important for systems control, in the present application in particular for the ability to detect the occurrence of own acts which were not as intended. Different approaches to predictive, quantitative models of detection have been taken.

A classic approach is based on the **signal detection theory** that was developed from studies of problems in statistical analysis of detection of radar signals in noise (Wald, 1947). A review of later developments can be found in Sheridan and Ferrell (1974). An extensive number of experiments have been made with subjects detecting visual or auditory signals in a background of noise. Given probability distributions of noise and signal plus noise, signal detection theory can lead to prediction of the probability of misses and false alarms in a detection task. The theory has, however, had more application in experimental psychophysics than in systems design.

Detection of changes in system behavior from signals, rather than appearance of signals in noise, has been modeled by **estimation theory**. This is a control theoretic approach based on parameter estimation in a state space representation. Models based on estimation theory represent human detectors as **ideal observers** monitoring dynamic processes. Gai and Curry (1976a) have proposed such a model which assumes that the human observer can be represented by an optimal Kalman filter which serves to eliminate perceptual noise and to predict the state of the system monitored. Detection is, then, based on the Kalman filter residuals, i.e., the difference between the observers' expectations regarding the states and the actually observed states. In this type of detection model, only visual input information is considered, i.e., the system is an open loop. It has been suggested, partly based on Young's (1969) experiments, that feedback from the hand movement of a closed-loop human controller will improve failure detection (Curry and Ephrath, 1976). Subsequent experiments tend to indicate that the influence on detection of closed-loop control depends on circumstances. A series of experiments made by Wickens and Kessel (1981) confirm the hypothesis of improved closed-loop detection, whereas Ephrath and Young (1981) report ambiguous results possibly depending on the level of workload in the task. Another explanation may be that perception and motor schema generation have different characteristics and limit properties in their functions of dynamic world modeling. Probably, comparative experiments with detection in open-loop monitoring and closed-loop control could be used to separate aspects of these functions.

The attention allocation and detection functions basically depend on signal processing. The information considered is the quantitative state of signal patterns related to spatially distributed sources with reference to their stochastic nature. For these functions, the functional meaning with respect to selection of proper action is not considered. This interpretation is represented separately in judgment- or decision-making models.

**Manual Control Models**

Models of human performance in closed-loop control tasks have been very important for representation of the properties and limitations of humans in vehicle control, in particular in aviation, and a separate school of modeling based on control theoretic tools has developed.

A review of models of human sensorimotor performance has been given by Pew (1974). Such models have been developed, in particular, from laboratory tracking tasks, and will typically give information on signal-to-noise ratio, maximum bandwidth when tracking unpredictable wave forms, prediction capabilities in sine-wave tracking, etc. The role of predictive feed-forward control for an industrial control task has been demonstrated experimentally by Crossman and Cooke (1962). For design and evaluation of the control system for, e.g., aircraft, models based on continuous linear differential equations are important, because they are suitable for computer simulation of the total system performance. A review of such models can be found in Sheridan and Ferrell, (1974).

The most recent development of such models has led to the optimal control models which are based on the observation of Leonard (1960) and Roig (1962) that the mean-square error from human tracking data approximated the mean-square error of various optimal controllers. Optimal control models have in particular been developed by Baron and Kleinman (1969) and used for describing pilot performance (Kleinman et al.1971). This model is based on the assumption that a well-trained, well-motivated human operator will act in a near-optimal way subject to certain internal constraints which limit his behavior. The internal dynamic world model necessary to account for human anticipation is represented by a Kalman-Bucy optimal filter. The model also includes observation noise and time delays depending on the instrument scanning strategy. The criterion is typically used to minimize square deviations from desired output as well as squared control effort according to a chosen trade-off ratio.

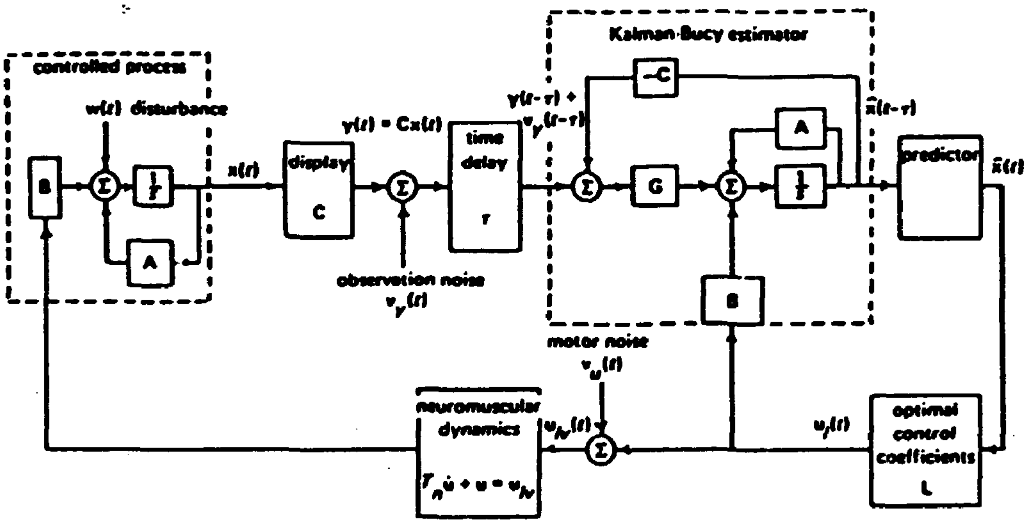



Figure 2. A typical optimal control model of a human vehicle controller. (From Kleinman, Baron, and Levison (1971) with permission from I.E.E.E.).

The model has been developed beyond the simple man-in-a-servoloop case, in that higher-level sequential functions for parameter and criterion control have been added in order to include multivariable control, monitoring, and decision making. This effort has proved successful for flight control and landing approach planning (Muralidharan and Baron, 1980; Baron et al., 1981). Efforts have also been made to extend this model to process control, and recently (Baron et al., 1982), a simulation model based on control theory has been proposed for simulation of the dynamic performance of a nuclear power plant including the operating staff. One important aspect of this approach is that human behavior at all three levels (skill-, rule-, and knowledge-based) as well as their interactions are considered in one integral model. It appears, however, at present to be difficult to collect the explicit human performance data that are needed for implementation of the model.

This kind of integrated model has great importance for design of aviation and other vehicle systems because the decision and manual control tasks of the operator form one integrated task in direct coupling with system dynamics. The task of a pilot or driver is a direct space-time control of a moving physical object, the vehicle. The sensorimotor level of the human information processing in this task serves for control signal processing; i.e., the output manual actions are continuous **signals** in the semiotic sense.

For process plants of the present levels of automation the continuous control **signal** processing is, however, automated. This means that human output actions will typically be related to switching and valving, and will be interpreted as stereotyped **signs** by the plant systems. This means that the human sensorimotor behavior will largely be occupied used for an interface manipulation skill. For this, the time-space characteristics will have no direct relation to the basic system dynamics or the supervisory control tasks for which they in a way act as separating interface. For such systems, the interface manipulation skill is probably most conveniently described separately with reference to the parametric description in terms of the gain-bandwidth-accuracy mentioned above. In particular, the state identification can only be represented in terms of a Kalman observer when the dynamic properties of the system to be controlled are known in terms of the parameters of control theoretic state space description. This is only the case for well-structured systems as, for instance, aircraft and space vehicles, and to some extent, the internal thermodynamic processes of industrial plants, and not for object manipulation in the physical environment in general. In consequence, models based on optimal control theory are only suited to represent the state identification of sensorimotor behavior, and the feature formation necessary to release and modify skilled patterns in case of manual control of well-structured dynamic systems, from which information is interpreted as **signals**. In less-structured situations when feature formation and state identification are based on recognition of information interpreted as **signs**, models of human judgment in terms of statistical representations are more suitable.

## Models of Human Judgment

State identification and recognition based on a pattern of cues or signs are typical front-end functions involved in the behavior at the rule-based level. Patterns of information from the environment - or the internal representation of the problem space - are interpreted as signs referring to manual acts or information processes. The process is not based on functional or symbolic reasoning, and the association from sign to response may be based on purely empirical evidence from prior trials or learned from a teacher. The person will typically not be able to identify the information from the environment acting as cues for his decisions, and very probably conscious introspection may only lead to invalid after-rationalization, because the basis of introspection is a process of a higher cognitive domain than involved in the rule-based behavior itself.

Models of judgment and decision making in everyday professional life at the rule-based level have very important implications for all trades, in particular for education and training and for the design of supporting tools. Therefore, models of this process have been the topic of extensive research in the area called human decision making, judgment, or choice, depending upon the paradigm of the approach. Several different approaches can be identified in this modeling effort, ranging from purely normative models based on economic concepts to predictive models derived from psychologic research. (For a detailed review see Hammond et al. (1980).

**Decision Theory**

This model is a mathematical model based on the expected utility theory developed by economists (Morgenstern) and mathematicians (von Neumann). The theory can be traced back to Bernouli's work on the worth of a decision determined by the probability of events and their associated utilities. Decision theory limits its interest to the single-system case, which involves one person without full knowledge of the task situation and without feedback about the effect of the decision. The approach focuses on decision making from a prescriptive point of view only. It is a logical structure for decisions and makes no claim that it represents or describes the information processing of human decision makers. The emphasis is not on what they do, but what they should do. Modern theorists (Keeney and Raiffa, 1976) emphasize the mathematical modeling of subjective probability and utility and promote the use of the theory to aid decision makers to achieve logical consistency.

Classical decision theory is, however, not useful as a model of human performance in real-life tasks, because it is a normative model. It will, therefore, not be considered in more detail in the present context. It may, however, be an important candidate to consider for computer implementation in decision support systems.

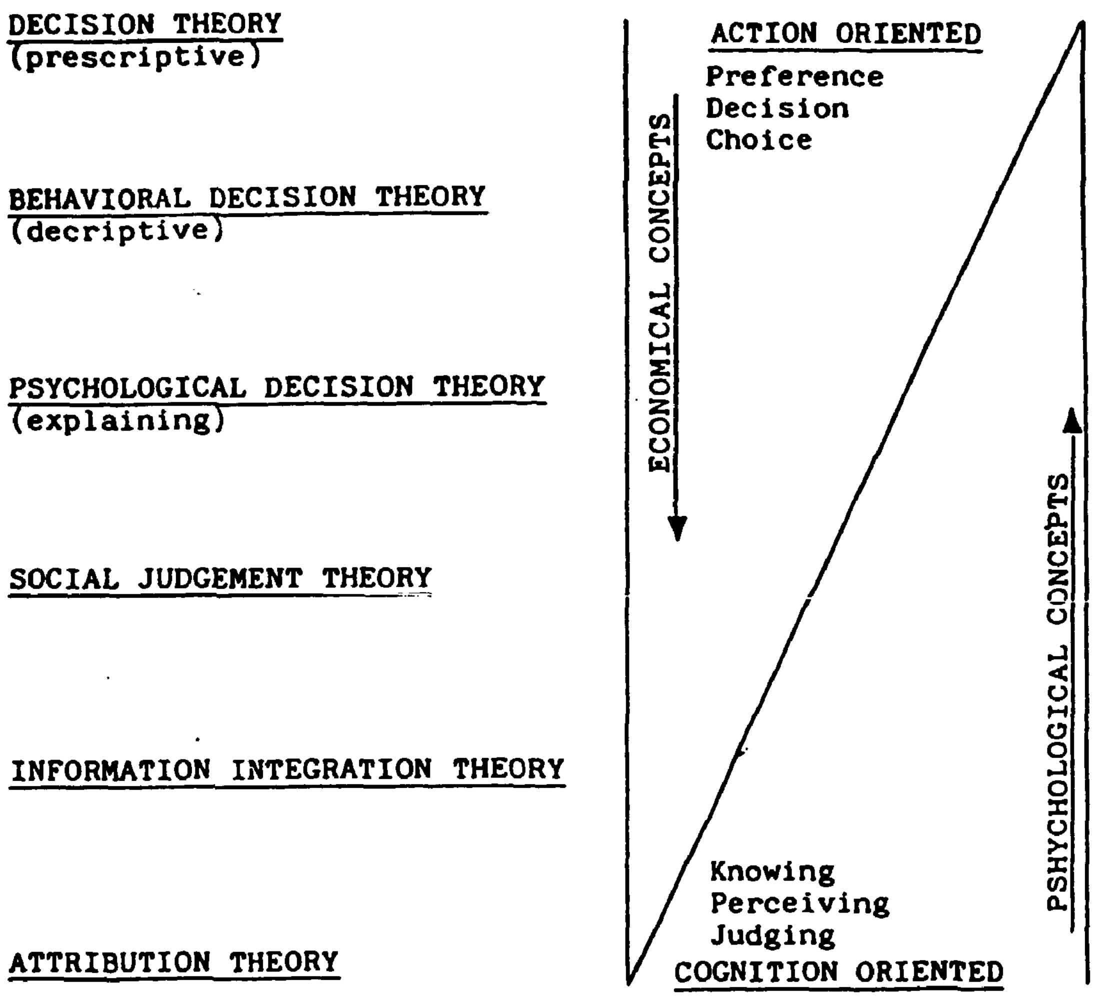



Figure 3. Different approaches to models of human decision and judgement and their relations to economical and psychological concepts. (Figures 44-46 have been adopted from Hammond et al. (1980) with permission from Hemisphere Publishing Corp.

## Behavioral Decision Theory

This theory was initially developed by Edwards and Tversky (1967) and is based on the Bayesian probability theory and on an experimental approach to modeling. The aim of the experiments has been to find out how closely human information processing approximates the Bayesian process, and how information is used to revise subjective probabilities. Experimental manipulations of the objective probabilities were used to examine revisions of the subjective probabilities.

The approach intends to describe human departures from optimal performance empirically and to explain such departures in terms of both the external (task) and the internal (psychologic) conditions. An often-mentioned departure from optimal decision making is conservatism, which represents the failure of humans to revise their posterior subjective probability as much as they should upon the receipt of new information. People are "conservative Bayesians".

Behavioral decision theory and classic decision theory share the same basic concepts as decision, preference, subjective probability, and utility, and both refer to the cognitive process of continuing probability and utility as "aggregation". These concepts are inherited from economic theory and mathematics, not psychology. However, unlike classic decision theory, behavioral decision theory does not aim at a prescription of the decision process, but at a description of the actual deviations from the optimal process. As such, it has had a rather extensive use to describe deviation from rational behavior in experiments on diagnostic behavior, for instance, in electronic troubleshooting (Rigney et al., 1968).

## Psychologic Decision Theory

In its recent form this theory is based on the work of Tversky and Kahneman (1974, 1979) with their concepts of representativeness, availability, and the use of heuristics.

Their approach rejects the use of optimal decisions as frame of reference for description. It turns instead to a search for the psychologic mechanisms that people use to evaluate frequencies and likelihoods. Psychological decision theorists generally consider subjective utility theory to be empirically unsuccessful and to be replaced by a theory that **explains** human behavior, that not only explains **why**, but is also able to predict **when** people replace the laws of statistics by heuristics. The theory is still rooted in decision theory, using probabilities and utilities as central descriptive forms, but focuses on the way in which people assign probabilities to events on explicit formulation of the **biases**, and upon classification of means for supporting decisions by **debiasing** because it appears that simply warning people against their bias often proves ineffective).

Tversky and Kahneman have identified a number of biasing heuristics which are relevant also for the human-machine context (Tversky and Kahneman, 1974). Because the basic role of the heuristics is "to reduce complex tasks to simpler judgmental operations" - they are relevant in many situations during disturbed plant operation.

The heuristics and their most important biasing effects are as follow:

* **Representativeness.** People associate to a prototyped number of a class, neglecting prior probabilities and base rate frequencies as well as the effect of sample sizes, and are influenced by the illusion of validity as well as by the misconception of regression toward the mean.

* **Availability.** People assess the frequency of a class or the probability of an event by the ease with which instances of their occurrence could be brought to mind. They are typically influereed by bias owing to easy retrieval of familiar and recent events, and by bias owing to the effectiveness of a search set and the imaginability of situations, independent of probability.

* **Adjustment and anchoring.** People make estimates an initial value which may be suggested by the formulation of a problem and which is generally not adequately adjusted.

The heuristics and bias must be viewed as characteristics of the cognitive process, not as the effects of emotional or motivational behavioral factors (as, e.g., wishful thinking). "The main cause for the failure to develop valid statistical intuitions is that events are normally not coded in terms of all the factors that are crucial to the learning of statistical rules" (Tversky and Kahneman, 1974), i.e., they operate from convenient signs, not defining attributes, as it is discussed below in considering errors of "familiar association shortcuts".

## Social Judgment Theory

This theoretical development as well as the basic concepts of "ecological validity of cues" and "utilization of cues" originate in the Brunswik's theory of perception. The primary intention of the approach is not to explain, but to describe human judgment processes, and to provide guides for the development of decision aids. Central to the development has been the work of Hammond, Brehmer, and others (see Hammond et al., 1980). The basic framework of the theory is Brunswik's "lens model" (see Figures 4&5). Central to the model, as shown in Figures 4&5, is that the environment and the person described symmetrical terms, hence, the name "lens model". The cues of the task vary in "ecological validity" and the person has a variation in "cue utilization", both of which may assume various linear and nonlinear forms. Both sides are analyzed and attempts are made to match ecological validities to cue utilization, as well as ecological function forms to subjective function forms, i.e., to identify the extent to which the principles of organization that control the task system are reflected in the principles of the organization that control the cognitive system of the person. The model, therefore, appear to be well suited as a basis for representing the effects of human error mechanisms, as will be discussed in detail in a later section.

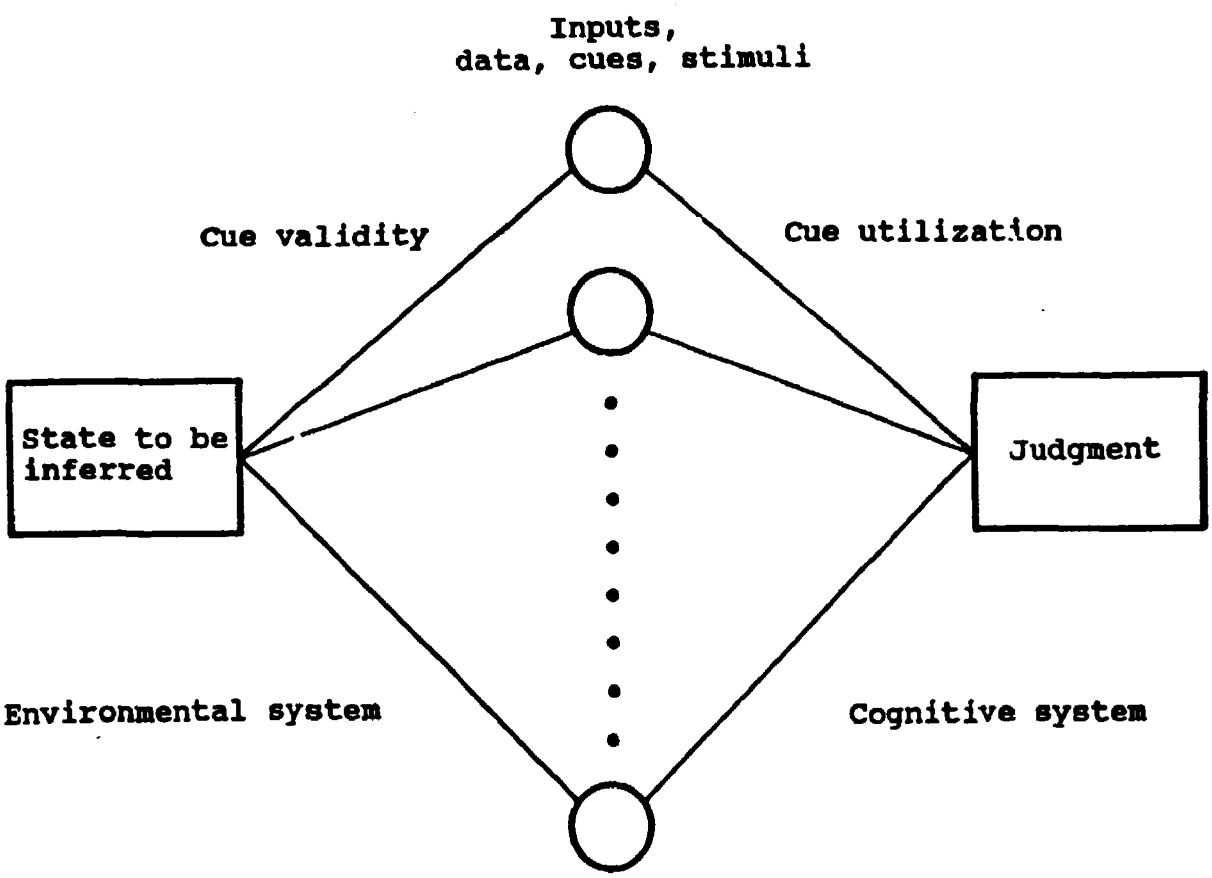
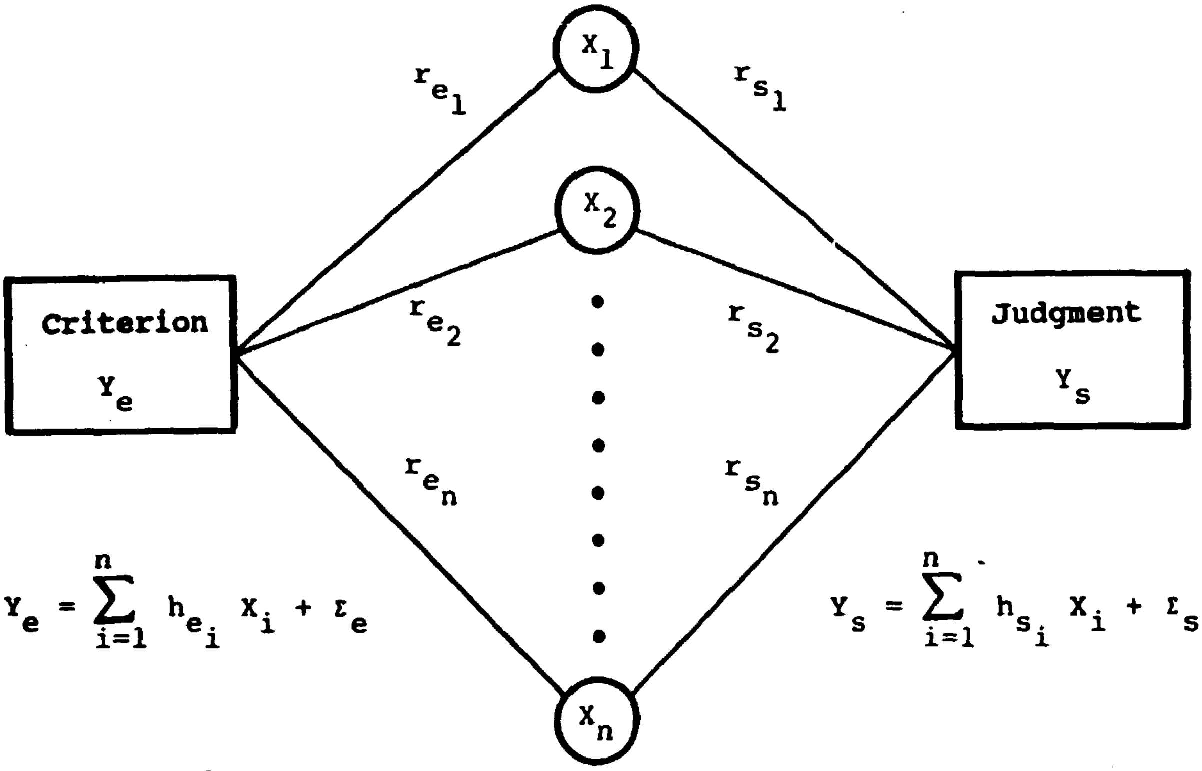



Figure 4 . Decision theories consider the cognitive system and the environmental system to varying degree.




$$Y_e = \sum_{i=1}^{n} h_{e_i} X_i + \varepsilon_e$$

$$Y_s = \sum_{i=1}^{n} h_{s_i} X_i + \varepsilon_s$$

$r_e$ = Ecological correlations

$r_s$ = Response correlations

$h_e, h_s$ = Optimum regression weights

<u>Figure 5</u>. Brunswik's "Lens Model".

The "lens model" has been the basis of research concerning diagnostic judgments in several professional activities such as, for instance, stockbrokers, clinical psychologists and physicians (Brehmer, 1981). A problem in such research is to describe a mental process of which the person is not himself aware. The approach has been to assume that even though the person is not aware of the process, he will know the information, i.e., the cues, on which the process is based. In experiments, therefore, cues identified as diagnostically relevant by expert judges are used to present, generally in written form, subjects with trial cases. From the experimental research then, the statistical model describing diagnostic behavior is identified. The general result has been that linear statistical models, such as multiple regression analysis, have been adequate. Four general results are typical of such diagnostic experiments. First, the judgment process tends to be very simple. Even though persons identify up to 10 cues to be relevant to diagnosis, they actually use very few, usually only two or three, and the process tends to be purely additive. Secondly, the process tends to be inconsistent. Subjects do not use the same rule from case to case, and judgment as a second presentation of a case may differ considerably from what it was the first time. Third, there are wide individual differences even among subjects with years of experience. They differ with respect to the cues used and the weights they apply. The fourth general result is that people are not very good at describing how they make judgments (Brehmer, 1981).

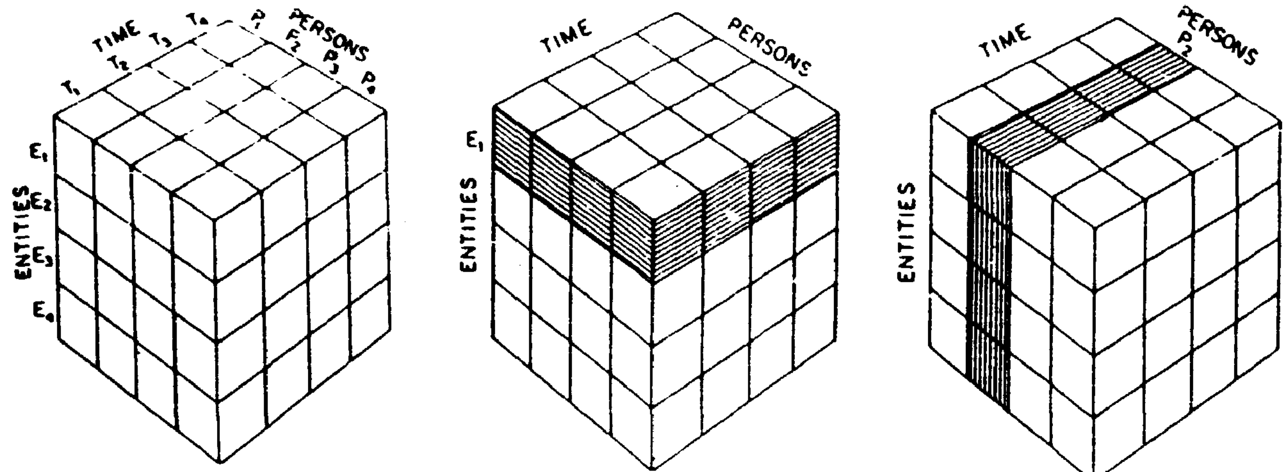
Given that intuitive judgments of an expert are performed by processes below the level of conscious attention, one may wonder whether laboratory experiments based on formal cues identified consciously by professional experts (textbook cues?) really correlate with the cues used in the real life situation in which cues depending upon prehistory and informal "situational cues" may play an important role.

## Attribution Theory

The basis of this theory is the theoretical work of Heider (1958) but has been developed by several researchers, for instance Kelly (1973), who has given several survey papers. The theory has been central within social psychology. The theory is very general in its aim and tends to cover not only a single person making judgment of the cause of an event in a social situation but also interpersonal relations in judgment, learning, and conflict. The central concept of the theory, "attribution", can be considered a special case of inference or judgment. Choice of action or preference for friends, for instance, follow from different attributions, but the theory is primarily concerned with inferences about causality, i.e., causal attributions, and should therefore probably be considered more in regard to human-machine systems than it presently is.

Kelly (1973) considers two different cases. First, the attributor has information from several observations and is able to respond to covariation between the observed effect and its possible causes. Second, the attributor has information from only a single observation and must respond to the set of conditions present at a given time. Thus, it is necessary for him to take account of the configuration of factors that are plausible causes. By the covariation concept, an effect is attributed to one of its possible causes with which it covaries over time. Implicit in this principle is the problem of the exact temporal relations assumed to exist between a cause and its effect.

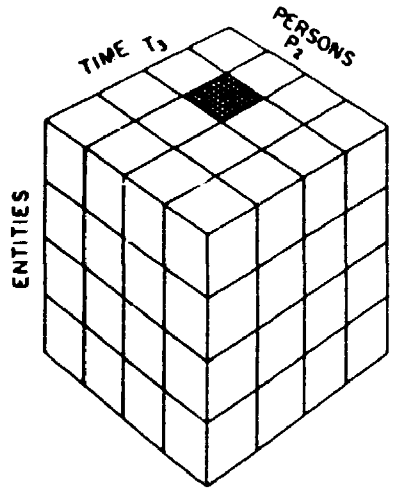
Kelly and Heider conceptualize the attribution process in terms of the analysis of variance as employed by the psychologist to interpret experimental results. "The assumption is that the man in the street, the naive psychologist, uses a naive version of the method used in science" (Kelly, 1973, p. 109). In this analysis of variance, the salient possible causes constitute the independent variables and the effect constitutes the dependent variable. For a wide range of attribution problems, the classes of possible causes are as shown in Figure 6 : persons, entities, times. An important application of the persons x entity




The analysis of variance framework for making causal inferences.

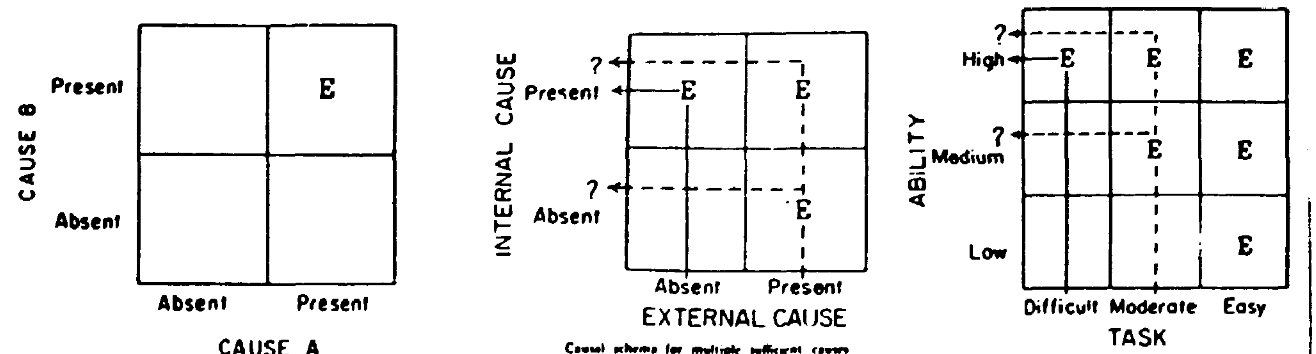
Data pattern indicating attribution to the entity

Data pattern indicating attribution to the person.




Data pattern indicating attribution to the "circumstances."

Analysis of variance of attribution theory. (Figures 47-48 have been reproduced from Kelley (1973) with permission from American Psychological Association).




Causal scheme for multiple necessary causes.

Causal schema for multiple sufficient causes

Causal schema for compensatory causes (E = success)

Figure 5 Causal schemas of attribution theory.

x time framework has to do with attribution validity. A response is known to be valid if (a) the response is distinctively associated with the stimulus, (b) there is consensus, other people have a similar response, and (c) the response is consistent over time. These criteria have been tested experimentally, and studies designed to test the idea that naive subjects treat information informally in a manner similar to the way statisticians treat it formally, have been supportive. Configuration concepts are the basis of causal inference on the basis of a single observation of the effect. A person is rarely acting in complete ignorance, in that ordinarily he has observed similar effects before and has some notions about possible relevant causes. Several statements can be made about the way attributors think: the discounting principle stating that a given cause in producing a given effect is discounted if other plausible causes are also present; the augmentation principle stating that if the external cause tends to inhibit or suppress the observed effect, the presence will increase the attribution of causes internal to the person observed, etc. The configuration concept has been implemented in a set of causal schemata related to compensatory causes, multiple necessary causes, and multiple sufficient causes, see Figure 6. Such causal schemata provide a person with means for making causal attributions given only limited information.

People have repertoires of causal schematas enabling them to deal with causal problems, and prototyped features can be identified. A major task for attribution theory is to specify when a given schema is evoked. In this function also stereotypes can be found: tendencies to prefer simple, e.g., single-cause patterns. This tendency has also been identified by Duncker (1945). He notes that certain cause-and-effect connections are intelligible to an attributor without evidence of covariation: a track resembles a foot of an animal; heavy things make loud noises, etc. Another simplification is that consequences will be linked directly to an actor rather than being viewed as a joint function of him and the situation.

As with other theories, systematic discrepancies between normative and intuitive inferences have been identified: subjects fail to extract all available information; they do not gain as much confidence from a series of events characterized by consistency as a probability model suggests they should, or they give too much weight to exceptions from general trends. A tendency to find causal explanations for variations that should be attributed to sampling variability is related to the bias from representativeness as discussed by Tversky and Kahneman (1974).

Attribution theory has, as mentioned, not been much considered in the context of human-machine systems but the results obtained should probably be considered, for instance, in relation to group decision and personnel management modeling and in relation to attribution of the cause of incident to personnel error. Finally, as control systems grow increasingly complex and functionally intransparent, the operator-control system relationship becomes more of an interpersonal relationship and, consequently, a social-psychologic approach such as attribution theory may be an appropriate point of view.

## Fuzzy Set Theory

In the typical models of human judgment, the relationship between the criterion and the attributes as well as between the attributes is considered to be stochastic in nature. This means that the categories and attributes are considered to be well defined and describable by classic logic including the rule of exclusion of the mean; i.e., an attribute is present or not present, and a judgment is either true or not true with a certain probability. The uncertainty may be due to the fact that the underlying phenomena are stochastic by nature, but it may also be due to lack of knowledge on part of the observer, to choice of attributes that are convenient cues rather than defining attributes, or may be caused by discrete representation in verbal statements of states of affairs which are in reality overlapping and fuzzy. This had led to extensive efforts to develop models of fuzzy reasoning (Gaines, 1976) based on Zadeh's fuzzy set theory

(Zadeh, 1965). An extensive bibliography on applications can be found in Gaines and Kohout (1977). The idea is basically that membership of a category is defined by a continuous membership function varying in the interval 0-1, for instance, the membership value of a temperature related to the category "hot" is a continuous function of measured temperature; the membership value of a man related to the class "tall" is a continuous function of his height around 180 cm. This means that the classic probability distributions are replaced by "possibility distributions". From the literature it seems that two different developments can be distinguished. One is the use of fuzzy sets in familiar many-valued logic; another is development of a fuzzy logic in which the truth values are themselves fuzzy sets. Doubts have been raised on the value of the latter approach (Haack, 1979) and, in general, the value of fuzzy sets compared with more traditional approaches seems to be a bit obscure.

Experimentally verified fuzzy models have typically been based on fuzzy sets that have been manipulated using nonfuzzy operations, which could have been handled by the traditional sum-of-weighted-attribute decision models. King and Mamdani (1977) have studied a fuzzy set model of a simple manual control task. Using verbal protocols they identified the following human operator control algorithms: if the error is "positive medium" or "big", and if its rate of change is "negative, small" then the control input should be "negative medium". The control action to choose from their observations was then selected from the possible actions as the one having the highest composite membership value. The individual membership grades were then quantified heuristically so that the performance of the algorithm was optimized. Comparison with performance of an algorithm based on conventional control design for control of practical systems proved that the fuzzy approach yielded better performance. A model based on a fuzzy set representation of human cement kiln operators' heuristic control rules has been developed by Umbers and King (1980) and afterward implemented in an automatic control system. They found it possible to obtain a satisfactory control, but experienced difficulties in having operators explain their behavior to a degree that made it possible to reach an adequate fuzzy model. The basic advantage of the fuzzy set approach seems to be its combatibility with ambiguous verbal statements.

Fuzzy set theory has been used by Hunt and Rouse to model the diagnostic behavior in fault finding tasks (Hunt and Rouse, 1981, 1983). They used a two-level model based on the distinction between symptomatic and topographical search during diagnostic tasks. The idea is that an expert will operate on the rule-based level by associating symptoms to proper repair acts as long as they find them useful. When this is no longer the case, they will turn to the knowledge-based topographical search based on the functional topology of the system. The definition of an expert then depends on the ability to realize when their expertise is no longer valid. This aspect will be discussed in more detail in relation to "Expert Systems" in the next section.

Hunt et al.(1983) define some symptomatic and topographical search rules that are derived from interviews of repair staff and laboratory experiments. The problem attacked by fuzzy set theory is the question which rule to apply in a given situation. In order to apply the knowledge contained in the rules, an algorithm for selecting rules is needed. Experience gained through prior diagnosis experiments has led Hunt and Rouse to the conclusion that four factors appear to significantly affect the rule selection process. For a rule to be selected: (a) the rule must be recalled; (b) the rule must be applicable to the current situation; (c) the rule must have some expected usefulness; and (d) the rule must be simple. It is hypothesized that rules are chosen on the basis of these four attributes. In most realistic situations, it is not always easy to make unambiguous assessments of the usefulness of a given rule. Further, from a human problem-solving point of view, recall and applicability are not simple either-or attributes. These attributes can, therefore, be considered to be fuzzy sets. In the model, the choice of the rule to apply in each instance is based on an evaluation of each of the available rules according to its membership value in the sets of recalled, applicable, useful and simple rules. The model has been tested experimentally, and the approach seems to be promising (Hunt and Rouse, 1982). An interesting feature of this model may be its potential for learning the membership functions by experience (Rouse and Hunt, 1981).

## Artificial Intelligence Models, Scripts and Plans, Expert Systems

During the recent decades there has been an extensive effort to design "artificial intelligence" systems based on analysis of human information-processing strategies. Different points of view concerning the aims of this development have been expressed varying from design of intelligent machines that consider only human behavior as a source of design ideas, to the view that AI programs are to be taken as models of human performance (Pringle, 1979). Taken with caution, i.e., viewing AI programs as one possible implementation of a cognitive model without necessarily an isomorphic relation between elementary processes of human and computer, the cognitive models of AI are at present the most promising approach to simulation of higher-level processes in system design.

Development within artificial intelligence has different typical periods and approaches; see, for instance, the review by Dreyfus (1981). The early period was mainly based on the general, context-free problem solving abilities of humans; a typical example is the general problem solver program (GPS) by Newell et al. (1956). This approach is tightly related to human behavior in the knowledge-based domain and is based on a representation of the basic causal structure of a problem, and will be discussed in more detail below. After some years of optimism, however, the development turned more toward models based on an extensive representation of domain-specific, procedural knowledge and of commonsense reasoning (Schank and Abelson, 1977). This means that the focus has shifted toward models of human behavior at the rule-based level, and typical for this approach are the various "expert systems" (Feigenbaum, 1979; Hayes-Roth et al., 1983). The knowledge representation in such systems is based on the "rules of thumb" used by human experts. These rules are not principles of general reasoning, and long inference chains that develop from general rules used on structural representations are rarely used by "expert systems". The basic deductions behind the rules of the real expert are not repeated; the systems have stored the symptom-based responses of an expert within a limited domain.

Recently there has been some concern about the reliability of such expert systems (Barnett, 1982). The symptom-based rules are derived from experts' experience rather than being model-based. Therefore, they are usually very effective and produce correct behavior, but they also have the potential to produce inconsistent responses to unfamiliar information. "The rules are plausible and work a high percentage of the time, this is why the expert uses them. However, when they fail, the human expert will know enough to realize this fact and find out why. He retreats to a better grounded model (one based upon more general principles)" (Barnett, 1982). In other words, the typical expert system only models human behavior at the rule-based level and lacks the ability to retreat to knowledge-based problem solving - it is only able to interpret information as signs. Therefore, the systems fail abruptly when the environment changes and no longer conforms with the experience behind the rules.

## Problem Solving Models, Artificial Intelligence Approaches

Problem solving at the knowledge-based level is, even now, best represented by the results of the classical analysis of verbal protocols by Selz (1922), Duncker (1942), de Groot (1963), etc. It is also typical that the approach taken by Simon et al. for simulating knowledge-based problem solving by means of production systems (Shaw, Newell and Simon, 1956; Newell and Simon, 1972) is based on this research. Based on a representation of the problem in a well-formed problem space, i.e., a model of the basic structure of the problem, behavior is developed from a set of production rules which are general, context-independent inference rules. A major weakness of this approach has been the difficulty of representing the ability of humans only to consider promising lines of reference for further development. This intuitive ability to see "where the problem is" is explicitly excluded from de Groot's analysis of the verbal protocols, and is not present in the production type models; see Dreyfus' comment on this issue (Dreyfus, 1972).

Generally, AI models have severe limitations because of the difficulty in representing the intuition or context present in human thinking stemming from the "subconscious world model" - or as Dreyfus expresses it - because of the lack of a physical body (Dreyfus, 1972). However, AI models are even then the best available tool for simulation of human information processing, and the development is important, not only for design of intelligent information-processing robots, but also for representing the human part in simulation of human-machine systems.

## Integrated Modelling Approaches

Recently, attempts have been made to develop integrated models which are suited for predicting operator responses to familiar as well as unfamiliar task requirements in complex real-life environments. Of special interest in the present context is the approach to models of supervisory control behaviour taken by Baron et al. (1982).

The characteristic feature of this approach is the closed-loop continuous modelling of the immediate interaction, based on the optimal control paradigm. This is an extension of the classical approach to vehicle control, and it has been successfully tested in this context. The "rule-based" addition to the optimal control model is based on production rules: IF (SITUATION), THEN (ACTION); or IF (SITUATION), THEN (SITUATION). This information is collected from procedure manuals or interviews with operating staff. If more than one situation is doomed to be "true" at a given time, choice is based on representation of perceived consequences. Closed-loop continuous behaviour maintains their own time scale, discrete actions use input data for completion times.

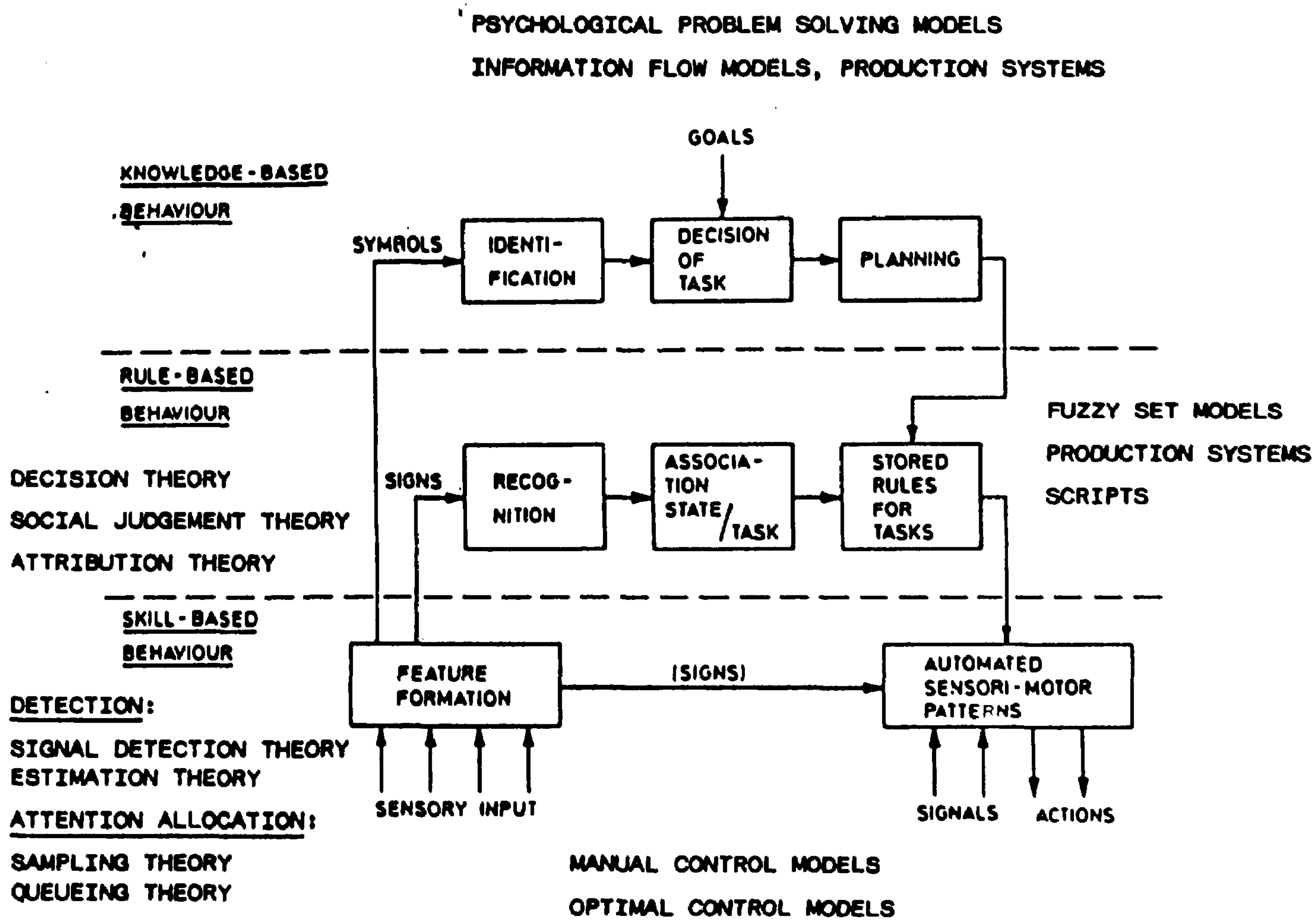



Figure 7. The various models of human behaviour are suitable for different functions of an overall model. An important problem is to model the interaction among the various domains of behaviour.

The fundamental operator functions included in the model are similar to the Ris| abstraction ladder: Monitoring, assessment (diagnosis), decision (choice of priority), implementation.

Basic elements of the model are:

Mathematical model of the plant environment:

1. State-space representation of the process including automatic control and engineered safety features.

2. A description of the potential disturbances (known and unknown).

3. A description of the instrumentation and display information supplied to the crew. (Note: in this respect the DYLAM may have an advantage being component based, the physical variables measured on the individual components may not be available from the functional balance simulation by a state space model). The model may be based on "first principles" or empirical models, e.g., records of data snap-shots.

The model of the human operator includes several functions:

1. Display processor. This function represents the "conscious" observation behaviour. A "single channel" processor is assumed. Visual fixation is allocated a certain time, visual observation noise is added. Auditory messages are stored, and the event detector notified. Chosen according to priority. The selection of a particular source of information is governed by goal-oriented processes, and will depend on the purpose for which information is being gathered. Much of the time the choice is supposed to be controlled by operating procedures, otherwise, a normative scanning strategy is assumed.

2. Information processor. Is based on algorithms for state estimation and prediction. Supplies "mental models" for fast-time predictions ("thought experiments") or evaluation of planned actions. The algorithm is similar to optimal Kalman filters. Can serve as event - detection based on residual between predicted and observed states. "Allows the modeling of all monitoring functions described by Kisner et al." is a stated claim.

3. Situation assessor. Several levels of assessment or diagnosis are contemplated. Lowest level, check of symptoms against templates residing in procedures or memory. Higher levels could include more sophisticated algorithms.

4. Response selector/formulator. Stored, formal procedures are triggered by the situation assessor. If not existing, informal procedures are formulated. How this is done is not settled: "They are included, partly in an attempt to model some knowledge-based behaviour, partly for modeling convenience". Decision making is required concerning the continuation and termination of activity, and selection among competing procedures. The choice is assumed to be rational and to be based on the computed "expected net gain".

5. Response effector. Three types of action: control, observation and communication. Associated with each action is an action time. When a procedure is activated, an operator is "locked in" for a certain time, and neither display observation nor other procedures will be considered. Exceptions are interrupts from high priority auditory information (alarms).

6. Memory. Is viewed as a store-house of information. Limitations of short term memory are considered.

7. Goal formation. In previous applications, a "unitary" goal structure was included. "In the present application probably a more hierarchical goal structure is required.

The structure of the model is, in fact, very similar to an integration of the various models described in the previous section, arranged in a hierarchical structure very similar to the skill-rule-knowledge framework. It appears doubtful, however, whether an attempt to build an integrated model will be worthwhile, until the information and data needed for the individual model domains, for instance the representation of the total task repertoire including the activating cue sets (templates) for a rule-based model, have been developed and tested.

Another approach to an integrated operator simulation model has been taken in the context of the DYLAM simulation system aiming at a simulation of the total course of events during incidents and accidents in nuclear power plants (Amendola, 1984). The line of development of a model simulating the cognitive processes of the operators (Mancini, 1986) leads to a structure similar to the above mentioned model proposed by Baron et al. (1982).

## REFERENCES

Barnett, J. A. (1982). Some issues of control in expert systems. Proceedings of the international Conference on Cybernetics and Society, Seattle, WA, October 28-30, 1982, 1-5.

Baron, S., and D. L. Kleinman. (1969). The Human as an optimal controller and information processor. IEEE Trans. Man-Mach. Syst. MSS-10, (1), 9-17.

Baron, S., G. Zacharias, R. Muralidharan, and R. Lancraft. (1981). PROCHU: A model for analyzing flight crew procedures in approach to landing. Proceedings of the Eighth World Congress of the International Federation of Automatic Control, Vol 7, 3481-3487. Oxford, England: Pergamon

Baron, S., C. Feehrer, R. Muralidharan, R. Pew, and P. Horwitz. (1982). An Approach to Modeling Supervisory Control of a Nuclear Power Plant. NUREG/CR-2988. ORNL/Sub/81-70523/1.

Brehmer, B. (1981). Models of diagnostic judgements, in J. Rasmussn and W.B. Rouse, eds. Human Detection and Diagnosis of System Failures. New York: Plenum Press.

Brehmer, B. (1984). Brunswickian psychology for the 1990s, in K. M. J. Lagerspetz and P. Niemi, eds. Psychology in the 1990s. New York: Elsevier Scientific Publishing Co.

Cacciabue, P. C. (1986). System Response Analyser: A Tool for the Analysis of the Behaviour of Systems. Tech. Note No. 1.05.C1.86.12 PER 1100/86. Ispra: Joint Research Center.

Carbonell, J. R. (1966). A queueing model of many-instrument visual sampling. IEEE trans. Hum. Fact. Electron. HFE-4 (4), 157-164.

Carbonell, J. R., J. L. Ward, and J. W. Senders (1968). A queueing model of visual sampling: Experimental validation. IEEE Trans. Man-Mach. Syst. MMS-9 (3), 82-87.

Dreyfus, H. L. (1972). What Computers Can't Do. New York: Harper and Row; New revised version, 1979.

Duncker, K. (1945). On problem solving. Psychological Monographs. Vol. 58, No. 5, Whole no. 270.

Edwards, W., and A. Tversky. (1976). Decision Making. Baltimore; Penquin Books.

Ephrath, A. R., and L. R. Young. (1981). Monitoring vs. man-in-loop detection or aircraft control failures, Human Detection and Diagnosis of System Failures, in J. Rasmussen and W.. B. Rouse, eds. New York: Plenum Press.

Feigenbaum, E. A. (1979). Themes and case studies of knowledge engineering, in Expert Systems in the Micro-Electronic Age, D. Michie, ed. Edinburgh University Press.

Gaines, B. R. (1976). Foundations of fuzzy reasoning. Int. J. Man-Mach. Stud., No 8, 623-668.

Gaines, B. R., and L. J. Kohout. (1977). The fuzzy decade: A bibliography of fyssy systems and closely related topics. Int. J. Man-Mach. Stud., (9), 1-68.

Haack, S. (1979). Do we need fuzzy logic? Int. J. Man-Mach. Stud., (11), 437-445.

Hammond, K. R., G. H. Mcclelland, and J. Mumpower. (1980). Human Judgment and Decision Making. New York: Hemisphere Publishing, Frederick A. Praeger.

Hayes-Roth, F., D. A. Waterman, and D. B. Lenat. (1983). Building Expert Systems. Reading, MA: Addison-Wesley Publishing Company.

Heider, F. (1958). The Psychology of Interpersonal Relations. New York: John Wiley & Sons.

Hunt, R. M., and W. B Rouse. (1984). A fuzzy rule-based model of human problem solving. IEEE Trans. Syst. Man. Cybern. Vol. SMC-14, (1), 112-120.

Keeney, R. L., and H. Raiffa. (1976). Decisions with Multiple Objectives, Preferences and Value Tradeoffs. New York: John Wiley & Sons.

Kelley, H. H. (1973). The processes of causal attribution. Am. Psychol. Vol 28, (2), 107-128.

King, P. J., and E. H. Mamdani. (1977). The application of fuzzy control systems to industrial processes. Automatica. 13, (3), 235-242.

Kleinman, D. L., S. Baron, and W. H. Levison. (1971). A control theoretic approach to manned-vehicle systems analysis. IEEE Trans. Automat. Control. AC-16, 824-832.

Leonard, T. (1960). Optimizing linear dynamics for human operated systems by minimizing the mean squared tracking error. WESCON, 4. (Pt. 4), 57-62.

Lind, M. (1982). Multilevel Flow Modeling of Process Plant for Diagnosis and Control. International Meeting on Thermal Nuclear Reactor Safety, Chicago. (Also: Roskilde, Denmark: Risø National Laboratory, Report No. M-2357).

Mancini, G. (1986): Modelling Humans and Machines. In: Hollnagel, E., Mancini, G., and Woods, D. (Eds.): Intelligent Decision Aids in Process Environments. Berlin: Springer Verlag, in press.

Muralidharan, R., and S. Baron. (1980). DEMON: A human operator model for decision making, monitoring and control. J. Cybern. Inf. Sci. (3), 97-122.

Newell, A., J. C. Shaw, and H. A. Simon. (1960). Report on a general problem-solving program for a computer. Information Processing: Proceedings of the International Conmference on Information Processing. Paris: UNESCO, 256-264.

Newell, A., and H. A. Simon. (1972). Human Problem Solving. Englewood Cliffs, NJ: Prentice Hall.

Pedersen, O. M. (1985). Human Risk Contributions in Industry: Guides for their Pre-Identification in Well-Structured Activities and for Post-Incident Analysis. Risø-M-2240. Roskilde: Risø Nat. Lab.

Pew, R. W. (1974). Human perceptual-motor performance, in Human Information Processing: Tutorials in Performance and Cognition. B. H. Kantowitz, ed. New York: Erlbaum.

Rasmussen, J., O. M. Pedersen, A. Mancini, A. Carnino, M. Griffon, and P. Gagnolet. (1981). Classification System for Reporting Events Involving Human Malfunction. Roskilde, Denmark: Risø National Laboratory, Report No. M-2240.

Rasmussen, J. (1986). On Information Processing and Human-Machine Interaction: An Approach to Cognitive Engineering. New York: Elsevier Scientific Publishing Co.

Reason, J. (1985). A framework for classifying errors, in J. Rasmussen, K. Duncan and J. Leplat, eds. New Technology and Human Error. London: Wiley and Sons. In press.

Reason, J. (1985). A preliminary classification of mistakes, in J. Rasmussen, K. Duncan and J. Leplat, eds. New Technology and Human Error. London: Wiley and Sons. In press.

Reason, J. (1985). The psychology of mistakes: A brief review of planning failures, in J. Rasmussen, K. Duncan and J. Leplat, eds. New Technology and Human Error. London: Wiley and Sons. In press.

Reason, J. (1985). Generic error modelling system (GEMS): A cognitive framework for locating common human error forms, in J. Rasmussen, K. Duncan and J. Leplat, eds. New Technology and Human Error. London: Wiley and Sons. In press.

Rigney, J. W., D. M. Towne, and A. K. Mason. (1968). An Analysis of Structure and Errros in Corrective Maintenance Work. TR-55, Electronics Personnel Research Group. Los Angeles: University Southern California.

Ringle, M. (1979). Philosophy and artificial intelligence, in Philosophical Perspectives in Artificial Intelligence. M. Ringle, ed. Sussex, England: The Harvester Press.

Roig, R. M. (1962). A comparison between human operator an optimum linear controller RMS-Error Performance. IRE Trans. Hum. Fact. Electron.. HFE-3 (1), 18-21.

Rouse, W. B. (1977). Human-computer interaction in multi-task situations. IEEE Trans. syst. man Cybern. SMC-7, (5), 384-392.

Rouse, W. B., and R. M. Hunt. (1981). A fuzzy rule-based model of human problem solving infault diagnosis tasks. Proceedings of the Eighth Triennial World Congress of the International Federation of Automatic Control. Kyoto, Japan, august 1981.

Schank, R. C., and R. P. Abelson. (1977). Scripts, Plans, Goals, and Understanding. Hillsdale, NJ: Lawrence Erlbaum & Assoc.

Selz, O. (1922). Zur Psychologie des produktiven Denkens und des Irrtums. Bonn: Friederich Cohen.

Senders, J. W. (1964). The Human Operator as a Monitor and Controller of Multi-Degree of Freedom Systems. ERE Trans. Hum. Fact. Electron . HFE-5, 2-5.

Senders,J. W. (1966). A reanalysis of pilot eye-movement data, IEEE Trans. Hum. Fact. Electron. HFE-7, 103-106.

Senders, J. W., and M. J. M. Posner. (1976). A queueing model of monitoring and supervisory behavior, in Monitoring Behavior and Supervisory Control. T. B. Sheridan and G. Johannsen, eds. New York: Plenum Press.

Sheridan, T B., and W. R. Ferrell. (1974). Man/Machine Systems: Information, Control, and Decision Models of Human Performance. Cambridge, MA: M.I.T. Press.

Simon, H. A. (1969). The Sciences of the Artificial. Cambridge, Ma: M.I.T. Press.

Tversky, A., and D. Kahneman. (1974). Judgment under uncertainty: Heuristics and biases. Science. 185, 1123-1124.

Umbers, I. G., and P. J. King. (1980). An analysis of Human Decision making in cement kiln control and the implications for automation. Int. J. Man-Mach. Stud. 12, 11-23.

Wald, A. (1947). Sequential Analysis. New York: John Wiley & Sons.

Wickens, C. D., and C. Kessel. (1981). Failure detection in dynamic systems, in Human Detection and Diagnosis of System Failures. J. Rasmussen and W. B. Rouse, eds. New York: Plenum Press.

Young, L. R. (1969). On adaptive manual control. Ergonomics 12, (4), 635-675.

Zadeh, L. A. (1965). Fuzzy sets. Inf. Control. 8, 338-353.

APPENDIX 2

**EXAMPLES OF THE HUMAN ERROR MODES OF THE TAXONOMY**

**USED FOR THE SIMULATION MODEL**

## SKILL-BASED LEVEL

### Motor variability:

* Inadequate precision leads to shortcircuit of terminals with screw driver.

* Inadequate precision in replacement of relay cover leads to short circuit of relay terminals.

* Varying use of force in manipulating a bank of valves occasionally leaves a valve leaking.

### Topographic misorientation:

* Failure in one of several pump trains in the basement leads to the decision in the control room to switch off the "north train"; however, during passage downstairs, an operator loses orientation and switches off the southern train, even though he has the proper intention.

**Stereotype takeover:**

* During normal operation of a process plant the power supply to the instrumentation is disrupted. Investigation reveals that the manual main circuit breaker in the power supply is in the off position. The conclusion was that a roving operator, in checking cooling towers and pumps, had inadvertently switched from a routine check round to the Friday afternoon shutdown check round and turned off the supply. The routes of the two checkrounds are the same, except that the opereator is supposed to pass by the door of the generator room on the routine check, but to enter and turn off the supply on the shutdown checks. Something en route has obviously conditioned him for shutdown checks (sunshine and daydreams?). The operator was not aware of his action, but did not reject the explanation.

* An experimental plant shuts down automatically during normal operation because of inadvertent manual operation of a cooling system shutoff valve. The valve control switch is placed behind the operating console, and so is the switch of a floodlight system, used for special operations monitored through closed-circuit television. The switches are neither similar nor closely positioned. The operator has to pass the valve switch on his way to the floodlight switch. In this case, the operator went behind the console to switch off the flood light, but operated the shutoff valves which caused plant shutdown through the interlock system.

* During start-up of a process plant, the plant is automatically shut down during manual adjustment of a cooling system. During start-up the operator monitored the temperature of the primary cooling system and controlled it by switching off and on the secondary cooling pumps to avoid water condensation in the primary system owing to the cold cooling water. On this occasion, he observed the temperature to pass the low limit, signaling a demand to switch off the secondary pumps, while he was talking to a cooperator about another matter over the phone. He then switched off the primary pumps and the plant immediately shut down automatically. He did not recognize the cause immediately,

but had to diagnose the situation from the warning signals. The control keys for the two sets of pumps are positioned far apart on the console. However, a special routine exists, during which the operator switches the primary pumps on and off to allow an operator in the basement to adjust pump valves after pump overhaul while they communicate by phone. Is the event caused by schema interference resulting from the phone call?

**RULE-BASED LEVEL**

**Forgetting an Isolated Item:**

* "Jumpers" not removed from terminals after repair;

* Switches not turned back to "operation" after instrument calibration;

* Bypass valves not reopened after pump repair;

* Cables not reconnected after instrument repair, etc.

**Incorrect Recall of Isolated Items:**

* Incorrect recall of numbers of valves and switches.

* Incorrect recall of figures, such as calibration references, set points, instrument readings.

**Mistake among alternatives**

* Using positive correction factors instead of negative;

* Using increasing, instead of decreasing signal in calibration.

* Disconnect pump A instead of B.

**Stereotype fixation**

* An operator presses air out of a plastic bag containing dust in order to seal it, although he knows it contains radioactive material. He gets contamination in his face.

* During a cleanup operation in a radioactive area, a vacuum cleaner fails. A foreman opens it for a possible rapid repair, despite the fact that he knows it contains radiosactive dust.

**Stereotype take-over**

* You have noticed that the road is icy and decided to drive carefully, but when a dog enters the road you kick the brake and .. .

* An operator enters an emergency procedure and executes a sequence of actions correctly but then inappropriately stops a pump, an act that follows the sequence in another, more frequently used procedure, but here is wrong and risky.

* An airplane is below acceptable altitude while approaching a runway. The pilot orders "full power" and the copilot responds correctly but also retracts the landing gear, resulting in a "wheels-up" landing. This act normally follows "full power" at low altitude during takeoff.

**Familiar association shortcut**

"The incident in question occurred in 1795, nine years after the discovery of the planet Uranus, and the principal figure involved was the great French astronomer Lalande. In that year Lalande failed to discover the planet Neptune, although the logic of events should have led him to it. Lalande was making a map of the heavens. Every night he would observe and record the stars in a small area, and on a following night would repeat the observations. Once, in a second mapping of a particular area, he found that the position of one star relative to others in that part of the map had shifted. Lalande was a good astronomer and knew that such a shift was unreasonable. He crossed out his first observation of the shifting point of light, put a question mark next to his second observation, and let the matter go. And so, not until half a century later did Neptune get added to the list of planets in the solar system. From the aberrant movement, Lalande might have made the inference not that an error had been made but that a new planet of the solar system was present. But he was reasonable. And it was more reasonable to infer that one had made an error in observation than that one had found a new planet."

The incident reveals that Lalande was not open-minded analyzing a physical system; he was absorbed in rule-based observation and data recording. (cited from Bruner et al., 1956):

* From a butadine explosion in Texas City; the investigation considers:

"Loss of butadine from the system through the leaking overhead line motor valve resulted in substantial changes in tray composition . .... The loss of liquid in the base of the column uncovered the calandria tubes, allowing the tube wall temperature to approach the temperature of the heat supply. The increased vinylacetylene concentration and high tube wall temperature set the stage for the explosion which followed. ...The make flow meter showed a continuous flow; however, the operator assumed that the meter was off calibration since the make motor valve was closed and the tracing of the chart was a straight line near the base of the chart. The column base level indicator showed a low level in the base of the column, but ample kettle vapor was being generated."

* An example from the melt-down of fuel elements in a nuclear reactor shows the great affinity to familiar signs, even a prewarning has been received:

"Certain tests required several hundred process coolant tubes to be blocked by neoprene disks. Seven disks were left in the system after the test, but were located by a test of the gauge system that monitors water presssure on each individual process tube. For some reason the gauge on one tube was overlooked, and it did not appear in a list of abnormal gauge readings prepared during the test. There was an additional opportunity to spot the blocked tube when a later test was performed on the system. This time the pressure for the tube definitely indicated a blocked tube. The shift supervisor failed, however, to recognize this indication of trouble. The gauge was adjusted at that time by an instrument mechanic to give a midscale reading which for that particular tube was false. This adjustment made it virtually certain that no flow condition would exist until serious damage resulted.

**Risø National Laboratory** **Risø - M -** 2616

| | |
|---|---|
| **Title and author(s)**<br>SIMULATION OF OPERATORS'<br>RESPONSE IN EMERGENCIES<br><br>Jens Rasmussen | **Date** September 1986 |
| | **Department or group**<br>Department of Computer & Information Science |
| | **Groups own registration number(s)**<br>R-7-86 |
| | **Project/contract no.** |
| **Pages** **Tables** **Illustrations** **References** | **ISBN** 87-550-1279-5 |

**Abstract (Max. 2000 char.)**

For the simulation of the accidental course of events in industrial process plants, a model is needed of operators' response to the cues presented by the system. A model is proposed, based on the simplifications which can be made when restricting attention to the operator functions having significants for a probabilistic risk analysis, and to only skill and rule based performance, i.e., to responses in the early phase of an accident. The model is based on Brunswik's lens model, a model of the normal task repertoire, and on a taxonomy of human errors.

To bring the model in perspective, a review of the state of the art of cognitive models of human behaviour is included.

**Descriptors**

Available on request from Risø Library, Risø National Laboratory, (Risø Bibliotek, Forskningscenter Risø),
P.O. Box 49, DK-4000 Roskilde, Denmark.
Telephone 02 37 12 12, ext. 2262. Telex: 43116, Telefax: 02 36 06 09